SEVENTH FRAMEWORK PROGRAMME

Inventory of existing PA-RIs cooperation – Update

FP7-ICT-248295/INFN/R/PU/D3.2 – PART I

Project Number: FP7-ICT-2007-4 248295

Project Title: Towards an Open and Sustainable ICT Research Infrastructure Strategy (OSIRIS)

Contractual Date of Deliverable: 30/11/2011 (month 23)

Actual Date of Delivery: **30/11/2011**

Workpackage contributing to the Deliverable: WP3 : Benchmarking

Nature of the Deliverable R

Dissemination level of Deliverable PU

Authors: INFN / Antonio Candiello, Mirco Mazzucato

Contributors:
IMCS / Baiba Kaškina [Sect. 2]
MFA / Géza Ódor [comments]
CSEM / Arno Hoogerwerf [hints on structure and models for ICT RIs]
ISERD / Felix Cincarevsky [bionano comments]

Reviewed by:
TNI-UCC / Alan Mathewson, Cian O'Murchu
CSIR / Colin Wright
IBBT / Piet Demeester, Rosette Vandenbroucke

Abstract:

This deliverable will list the most relevant cases of current cooperation between Public Authorities and ICT Research Infrastructures. This will include a standardized and structured high-level list of relevant ICT Research Infrastructures, which identifies the appropriate network RIs, computing RIs, Micro & Nano Technologies (MNT) and instrumental-related infrastructures, data infrastructures and Future Internet.

Keyword list:

Research Infrastructures, ICT, inventory, cooperation.

Clarification

Nature of the Deliverable

R	Report
P	Prototype
D	Demonstrator
O	Other

Dissemination level of Deliverable:

PU	Public
PP	Restricted to other programme participants (including the Commission Services)
RE	Restricted to a group specified by the consortium (including the Commission Services)
CO	Confidential, only for members of the consortium (including the Commission Services)

Disclaimer

The information, documentation and figures available in this deliverable, is written by the OSIRIS ("Towards an Open and Sustainable ICT Research Infrastructure Strategy"*) – project consortium under EC co-financing contract FP7-ICT-248295 and does not necessarily reflect the views of the European Commission.*

Executive summary

This document lists the most relevant cases of current cooperation between *Public Authorities* (PA) and *National Champions* (NC) within existing ICT *Research Infrastructures* (RIs) for the requirements of the Osiris project.

A high-level list of relevant ICT Research Infrastructure models is reported, which identifies the appropriate network RIs, computing RIs, *Micro & Nano Technologies* (MNT) and instrumental-related infrastructures, digital libraries frameworks. The Future Internet vision is also considered in this inventory.

The domains have been identified on the basis of several publications, including "The Future of Internet, Report from the National ICT Research Directors WGFI" [WGFI, 2008], "e-IRG Report on Data Management [DMTF, 2009], "Trends in European Research Infrastructures Analysis of data from the 2006/07 survey", [EC-ESF, 2007], "EGI Blueprint" [EGI_DS, 2008], "Riding the wave – How Europe can gain from the rising tide of scientific data" [HlEGSD, 2010] and the recommendations by consortium members and the stakeholder group.

For each domain, the relevant projects, the governance models and challenges experienced have been analysed. The consortium members provided the required input in order to better evaluate the cases of PA/NC cooperation within ICT RIs. A detailed list of intra-domain PA/NC-RIs collaborations is also reported.

This inventory will serve as a basis for the planned subsequent benchmark analysis.

Content

1. Introduction

This document lists the most relevant cases of current cooperation between *Public Authorities* (PAs) and *National Champions* (NCs) within existing ICT *Research Infrastructures* (RIs). PAs are National, supra-national, sub-national Governments or similar Public Bodies; NCs are Research Institutions (e.g., industry leader in a specific research or public area, National Research Centre) that drive or have exceptional knowledge in the chosen research field. Our task in this document is to analyse current examples of Public Authorities and National Champions cooperation and the practises employed regarding existing ICT RIs.

The consortium identified a set of key ICT RI *models*. Correspondingly have been identified the domains, reported in this Section, on the basis of several publications, including “The Future of Internet, Report from the National ICT Research Directors WGFI” [WGFI, 2008], “e-IRG Report on Data Management [DMTF, 2009], “Riding the wave – How Europe can gain from the rising tide of scientific data” [HlEGSD, 2010], “Trends in European Research Infrastructures Analysis of data from the 2006/07 survey”, [EC-ESF, 2007], “EGI Blueprint” [EGI_DS, 2008], and recommendations by consortium members and the stakeholder group.

Section 2 is mainly devoted to reporting on the Work Package (WP) #2 results relevant for the selection and prioritization of PA/NC-RIs cooperation cases. The challenges analyzed (see §2.1), and the results of the survey (see §2.2), both conducted in WP2, were taken into account in the inventory process in order to evaluate their specific relevance.

Consortium members have provided the required input to evaluate the cases of PA/NC cooperation within ICT RIs. Detailed analysis of each context lead us to structure an inventory of PA/NC-RIs *collaborations*, and this is the subject of Section 3.

Some final considerations will conclude the document.

1.1 Structure and Models of ICT RIs

As an intermediate step for an inventory of the existing ICT RI models of cooperation between Public Authorities and National Champions, a thorough analysis of the different ICT RIs has led the project to the conclusion that these have an intrinsic (topological and physical) structure that is built around the following elementary objects:

- *networks* – technological architectures that allow the high speed interconnection of different ICT facilities;
- *facilities* – sites with concentrated resources, such as *computing* facilities (as in HPC centres or smaller DCI computing sites/nodes) or *storage* facilities as in scientific repositories, digital libraries and DCI storage sites/nodes;
- *instruments* – scientific equipment capable of creating (in particular, digital) data from experiments; these provide new scientific data sets or repositories;
- *testbeds* – artificial environments needed to conceptualise, set up and test new kinds of ICT/internet interactions between humans and things;
- *laboratories* – plants for test and production of new ICT physical devices.

Networks are necessary to inter-connect the facilities at high speed – a prerequisite for distributed infrastructures. They also allow the connection of instruments, testbeds and laboratories to each other. Because the network grid is a minimum requirement for the cooperation between other facilities, its need is unquestioned. In general, it is implemented with top-of-the-shelf equipment and there is in general only one provider of the research network infrastructure in each country. *Facilities* can be computing facilities or storage facilities. *Computing* facilities generally work together in the distributed form of *Distributed Computing Infrastructures* (DCI) or in the concentrated form of *High Performance Computing* (HPC) facilities. Even HPC facilities can be networked so as to operate synergistically. The computing facilities gain from the cooperation in DCI and HPC through load optimization. Available resources that suit given requirements may be automatically selected at the user’s request. This process is ruled by agreements between the stakeholders. Corresponding middleware to implement this process is a requirement. *Storage* facilities store the growing amounts of scientific data. These storage facilities gain from cooperation because it allows them to distribute data over different centres, achieve redundancy and interoperability and also to back-up all the data. Much effort is needed to make this data accessible, this starts by achieving a uniform metadata. A lot of issues concerning data have to be tackled including data ownership, responsibility for the content, IPR, security.

Instruments are facilities that help to improve our understanding of complex problems. They frequently generate much data for later analysis. Typically, for each type of problem, there is only one, or a very small number of instruments in which many parties participate to share the

[14] In [PRACE D2.1.1, 2008] a detailed benchmark between different legal entities was done.

cost. *Testbeds* are needed to test and verify newly developed standards. The size and scale of these testbeds will have to be non-negligible to be able to represent real-life situations, especially for the Future Internet.

Laboratories are less tightly coupled than other ICT RIs. They mainly *produce* ICT in the form of physical devices with new technologies. Laboratories are not currently heavy consumers of ICT RIs – they do not produce much (digital) data, and in general they are not necessarily connected to the grid, because they do not have a particular need for access to data or computing power. Their outputs do, in general, have a direct economic impact in the short to medium term. As a result, laboratories often compete to obtain results first. Only recently, the increasing cost and complexity of the laboratory RI has forced them to start to cooperate in order to become more complementary, since no single RI can cover the entire field of ICT technologies anymore.

1.2 Inventory of ICT RIs Environments

The five elementary objects identified helped us to clarify the intrinsic nature of the ICT RIs as follows:

- *networks* are the structural element underpinning the *Research Network Environment*; they also create the structure for the *Research DCI framework*, as DCIs are based on a hybrid model that mixes computing, network and storage in order to obtain the best efficiency compromises;
- *facilities* are mainly a characteristic of HPC; at a smaller scale, they are constitutive elements also for DCIs and for *Research Data Infrastructures*;
- *instruments* are the target of the *Remote Instruments access model*, which is an ICT interface to instruments;
- *testbeds* are required in order to outline the *Future Internet vision*, which is first and foremost a standardization effort;
- *laboratories* are the principal objects around which the main structure of the *MNT collaboration framework* is built; the MNT model however also has other interesting features (interconnection, complementarity / no overlapping, integration, aggregation of capabilities).

We complete our analysis by selecting the following **ICT RIs environments** as the main reference points to identify relevant cases of current cooperation between Public Authorities and National Champions within existing and future ICT RIs, namely:

(E1) European and National *Network* RI environment,

(E2) (Grid & Cloud) European and National Research DCI framework,

(E3) (HPC) *high end / performance parallel computing* RI ecosystem,

(E4) MNT collaboration *facilities interchange* RI framework,

(E5) *Research Data Infrastructure* framework,

(E6) *Remote Instruments* access model,

(E7) *Future Internet* (FI) service-oriented vision.

In the next paragraphs, a brief description of each will be given.

1.3 The European and National Network RI Environment

The high bandwidth European Research Network provides to the research sector the connectivity over which other pan-European RIs can be put into operation. The network service use case is simplest for this group and its governance model the oldest – GÉANT is currently celebrating ten years of operation.

GÉANT provides Europe with a gateway for global collaboration, enabling researchers, students, teachers and other staff in institutions across the continent to participate in joint projects with their peers in other parts of the world. GÉANT offers different connectivity options and supporting network services to institutions, so enabling them to support the networking needs of their project and individuals.

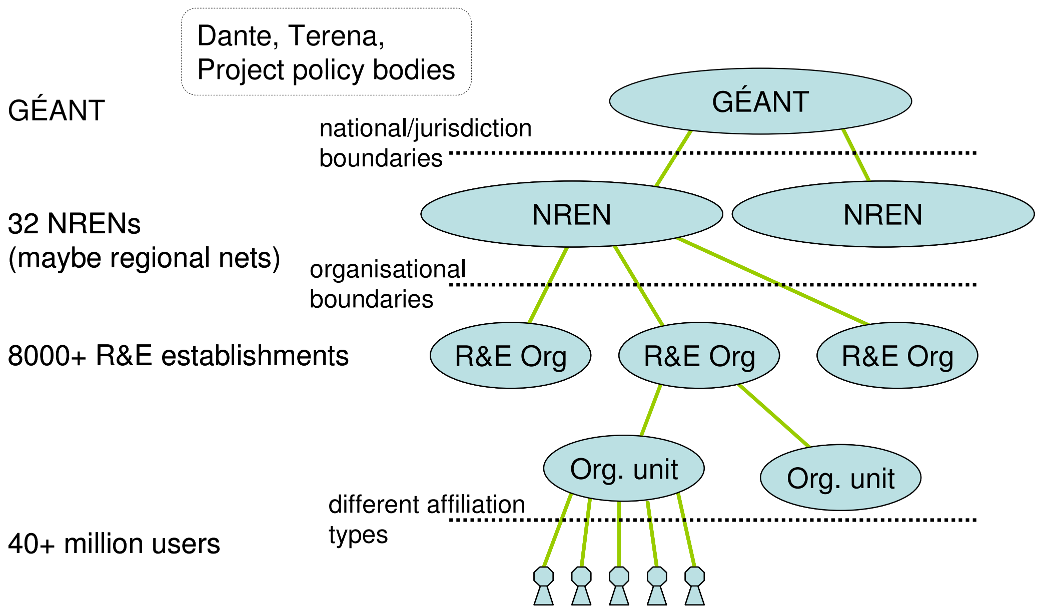

Fig. 1 – The Network RI Governance Model (image modified from a 2007 SWITCH NREN presentation).

The mature governance model of Network RIs is structured on the per-country specialized network service providers, namely *National Research and Education Networks* (NRENs), supporting the needs of the research and education communities. The NREN Policy Committee, with one representative per country, provides the general policy and coordination body at European level. The GÉANT series of projects (currently GN3) are co-funded by the NRENs and the EU to guarantee harmonized necessary innovation and future evolution. GÉANT is coordinated by the managing partner, namely *Delivery of Advanced Network Technology to Europe* (DANTE), which operates the European general backbone, and is supported by the dissemination and research partner namely *Trans-European Research and Education Networking Association* (TERENA), which provides a Forum for the coordination of the R&D efforts between the NRENs.

Other pan-European ICT RIs have utilised the NRENs' governance model as a reference.

As NRENs mature, their focus is moving increasingly towards added value services which become feasible by virtue of the existence of NRENs, for example: Eduroam enabling researcher mobility, or central video services, enabling global research collaboration.

International network collaborations are realised (see the following Fig.1) with North America (DICE), South America (ALICE2/RedCLARA), Africa (EUMEDCONNECT, UbuntuNet Alliance), India, South East Asia (TEIN3, CAREN, BSI).

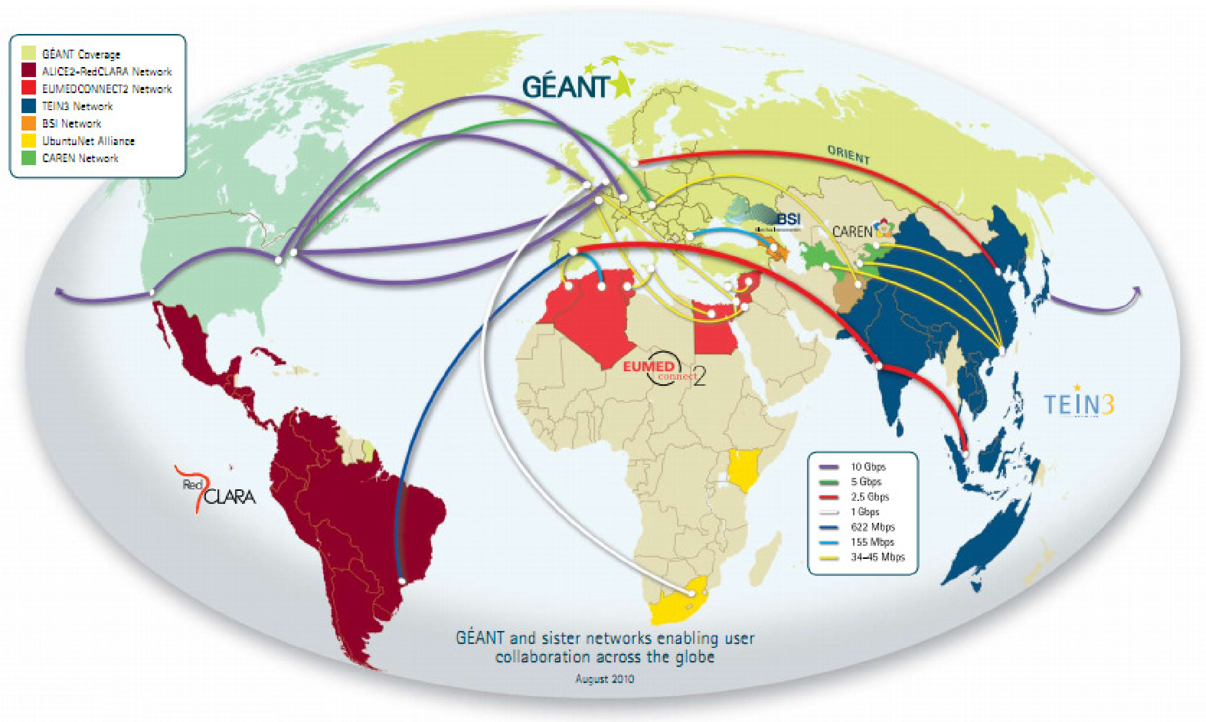

Fig. 2 – The GÉANT Network Environment and its worldwide interconnections (image from www.geant.net).

1.4 The European and National Research DCI Framework

1.4.1 Research Distributed Compute Infrastructures (Grids & Cloud)

The pan-European *Distributed Computing (& Storage) Infrastructure* (DCI), incorporated as *European Grid Initiative* (EGI), includes more than 300 Compute and Storage Resource Centres in Europe to which it arranges general access and sharing of services.EGI consolidates ten years of research and development that achieved by the *Enabling Grids for E-sciencE* (EGEE) series of European projects.

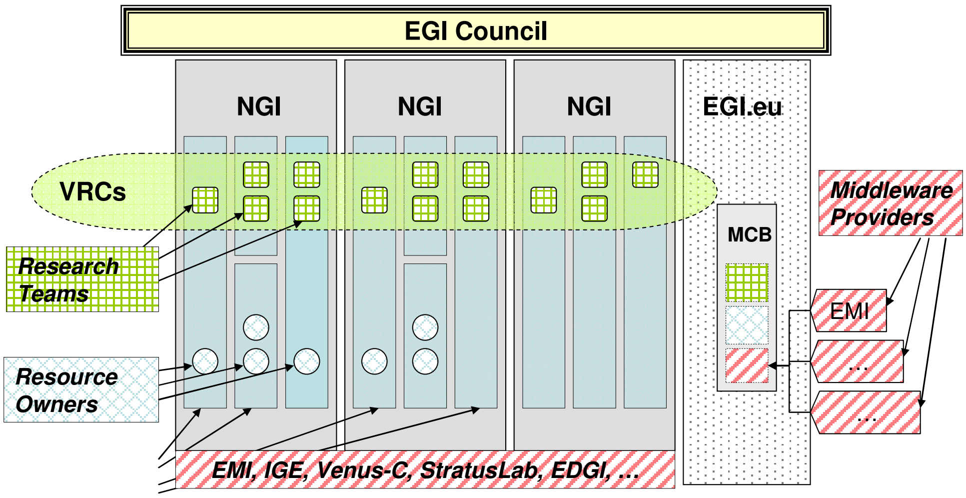

Fig. 3 – The EGI Governance Model (image modified from EGI_DS documents).

The EGI infrastructure is connected world-wide through *regional collaborations* with Latin America (current and past projects ACTION-GRID, GISELA, EELA-2, and EELA), Asia (current and past projects EUAsiaGrid, EUIndiaGrid, EUChinaGrid), Mediterranean and South East Europe (EUMedGrid, BALTICGRID-II, Baltic Grid, SEE-GRID-SCI, SEE-GRID2, SEE-GRID). Other projects are working to consolidate the international collaborations. The *Co-ordination and Harmonisation of Advanced e-Infrastructures* (CHAIN) is aiming at a harmonised and optimised interaction model for e-Infrastructure interfaces between Europe and the rest of the world. The *Exchange Programme to advanced e-Infrastructure Know-How* (EPIKH) is aiming to reinforce the impact of e-Infrastructures in scientific research and broaden the engagement in e-Science activities / collaborations both geographically and across disciplines.

EGI.eu, a non-profit organisation under Dutch law and located in Amsterdam, manages the European grid infrastructure. Its governance structure is partially based on a modification of the NRENs' model that can accommodate the added complexity of middleware R&D and user community management, but has a different funding model. The EGI governance model was designed according to the framework of the EGI Design Study project and then later refined, is based:

a) on the per-country *National Grid Initiative* (NGIs), responsible for the national grid e-Infrastructure and for maintaining relationships with customers,

b) by the EGI.eu organization which provides global services to all NGIs, and

c) by the EGI Council which provides the general coordination and policy decision framework.

The EGI *Integrated Sustainable Pan-European Infrastructure for Researchers in Europe* (EGI-InSPIRE) project has as goal to guarantee the overall operation of the European DCI together with its innovation and harmonized evolution and is co-funded by the NGIs and the EU.

The necessary *middleware* was not available from commercial providers at the time of the creation of scientific DCIs, so several years of research and development were needed to provide a production-class middleware layer. Contrary to the Networking RI, which is based on commercial products, the DCIs are based on software services provided by middleware providers coming, in general, from the Research sector. Currently, the EGI community has signed agreements with the *European Middleware Initiative* (EMI) – a close collaboration between major middleware providers, ARC, gLite, UNICORE, and specialised software providers like dCache – and the *Initiative for Globus in Europe* (IGE). Other middleware providers are EDGI for desktop grids, Venus-C and Stratus Lab for the cloud environment. All the middleware providers aim at supporting and evolving the EGI model with activities of research and development in the software domain.

As key stakeholders, the *Virtual Research Communities* (VRCs) are managed in EGI via specific support processes caring for them and for new user communities' development.

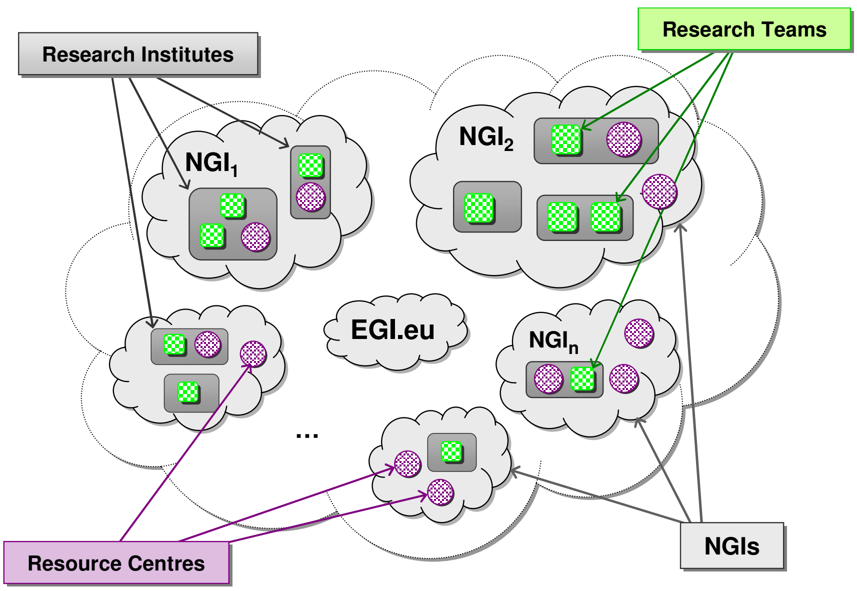

Fig. 4 – EGI Stakeholders. Research Teams are organized in thematic VRCs (image modified from EGI_DS documents).

1.4.2 The service model of grid & cloud DCIs

Scientific computing grids generally adopt a *shared use* model, namely that user communities share distributed computing and storage resources and obtain shared access to grid resources. The implicit extensibility of this model to new user communities and its intrinsic sustainability explains the popularity of grid computing in the research community. The counterpart is that users (and site managers) have to manage the added complexity of the middleware layer.

The *cloud paradigm* introduces further flexibility and elasticity in the DCIs offering and has favoured the decoupling of the physical layer from the software environments available to users at the same time. On the other hand, it favours a rapidly growing service model that is also being warmly accepted in the private industry sector due to its simplicity. The introduction of cloud services in DCIs will allow research teams to acquire the resources, or even better the software environments they need, in a much simpler way, together with the possibility of paying for what they use. Clouds offer users scalable and elastic access to computing resources while at he same time leaving complete freedom of the software environment choice, bringing closer effective *Infrastructure as a Service* (IaaS). The EU StratusLab project builds such a service. International policy, dissemination, technical standardization and grid-cloud interoperability are coordinated via specific initiatives, notably the *Standards and Interoperability for eInfrastructure Implementation Initiative* (SIENA), that aims to "define scenarios, identify trends, investigate the innovation and impact sparked by cloud and grid computing, and deliver insight into how standards and the policy framework is defining and shaping current and future development and deployment in Europe and globally".

1.5 The High End/Performance Parallel Computing RI Ecosystem

The European *High Performance Computing* (HPC) infrastructure is targeted towards providing the highest sustained parallel peak computing capacities to scientific research communities. The Partnership for Advanced Computing in Europe (PRACE), which has already commenced operation will maintain a pan-European HPC computing infrastructure consisting of up to six *top of the line* leadership systems (Tier-0) which are well integrated into the European HPC ecosystem as a single European entity. Each system will provide computing power of several Petaflop/s, in the longer term targeting the Exaflop/s computing power range. Four nations (France, Germany, Italy and Spain) have agreed to collectively provide 400 million Euro to implement supercomputers with a combined computing power in the multi Petaflop/s range over the next five years. This funding is complemented by up to 70 million Euros from the European Commission which is supporting the preparation and implementation of this infrastructure.

As is common in the supercomputing field, the PRACE service model requires that researchers from across Europe apply for HPC compute-time from a series of hosting nations via a central peer review process. Calls for proposals for computer time on PRACE RI machines are issued regularly and have strict closing dates. Computing resources in the HPC use case are scarce, so user communities have to compete against each other in order to gain access to them. The pan-European peer review system for PRACE is laid out in the statutes. It is based on the principle: "the best systems for the best science" and takes into account the results of the *HPC in Europe Taskforce* (HET) group and an analysis of national HPC peer review systems. Throughout the [PRACE D2.4.2, 2008] deliverable, the eight principles for PRACE peer review (*Transparency*, *Expert Assessment*, *Confidentiality*, *Prioritisation*, *Right to Reply*, *Managing Interests*, *No Parallel Assessment*, *Ensure Fairness to the Science Proposed*) have been considered.

PRACE is also seen as an enhancement of the *Distributed European Infrastructure for Supercomputing Applications* (DEISA) series of projects, targeted at the consolidation of a persistent European HPC ecosystem. The DEISA consortium has deployed middleware that enables transparent access to distributed resources, high performance data sharing at European scale, and transparent job migration across similar platforms.

On June 9^{th} 2010, the non-profit international association of nineteen members, named "Partnership, for Advanced Computing in Europe ASBL"[4], was set up in Brussels to represent PRACE; the PRACE Council held a constitutional meeting, at which it elected the Council Chair and the *Board of Directors* (BoD) Chair. An SSC Establishment Workshop, organised by six independent representatives of the European scientific communities then nominated the candidates for the PRACE *Scientific Steering Committee* (SSC). With the Council, the BoD, and the SSC in place, PRACE has implemented the main bodies of the governance structure that were developed in the project.

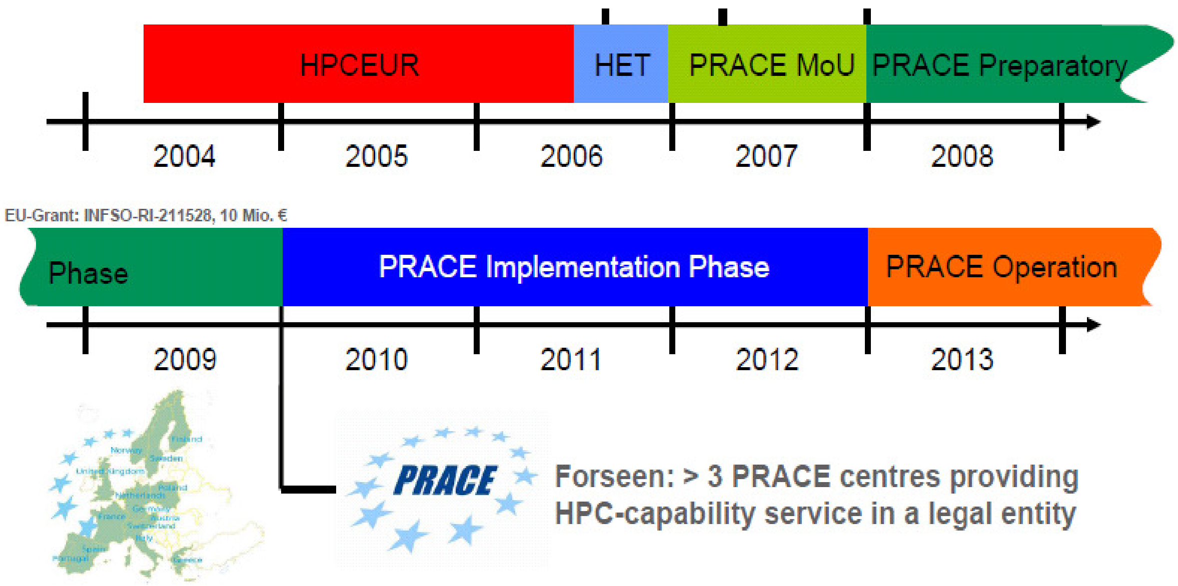

Fig. 5 – The PRACE Roadmap.

The four Principal Partner countries (called hosting members in the PRACE AISBL), namely Germany, France, Spain and Italy, have already made binding commitments to contribute Tier-0 resources, each with a value of € 100 Million in the next five years; the Netherlands followed subsequently in 2010. The ASBL currently has twenty members representing the fourteen countries that participated in the PRACE project plus six additional countries (Bulgaria, Cyprus, Czech Republic, Ireland, Serbia, and Turkey) that have joined the PRACE Initiative since the project started in 2008.

Fig. 6 – The PRACE/DEISA (and EGI) performance pyramid.

It has to be noted that *PRACE facilitated the transition from competition to cooperation mode amongst European HPC centres, which is an achievement per se*. It has led to *an exchange of best practice and expertise* in all aspects of HPC service provisioning among leading European centres, ranging from *procurement*, *infrastructure*, *operation* and *management*, *user support*, and *peer review* to *algorithms*, *code development* and HPC *technology*. These have been documented and serve not only as a model for the PRACE Tier-0 centres, but also for new national and focussed centres – especially in the new member states.

Regarding industry collaborations, PRACE has raised significant interest amongst European and international HPC vendors and technology providers. This has been leveraged to create a framework for information exchange and cooperation between PRACE and the HPC industry through the *PRACE advisory group for STRAtegic TechnolOgieS* (STRATOS). This structure will eventually foster the further development of a European HPC industry.

The open Grand Challenge problems of our time are so challenging that they will require properly researched and designed Exascale systems to make them tractable. Exascale designs will need to address the challenge of massive computational demands yet take into account the green imperative, the data tidal wave, current interconnect constraints, managing cooling and power demands, etc. The PRACE lessons will stand these developments in good stead.

1.6 The MNT Collaboration Facilities Interchange RI Framework

Considered as a research infrastructure, *Micro & Nano Technology* (MNT) is built around a limited number of extremely expensive facilities and laboratories, each requiring heavy investment and highly specialized personnel to maintain them. Due to its large fabrication facilities and clean rooms requirements, *microelectronics* is one of the most capital intensive production and research activities globally and is making contributions in the nanotechnology domain. At the present time *nanotechnology* requires more moderate requirements for its facilities, but still needs costly instrumentation and laboratories. Both of these will be considered in the following, but it should be noted that the dimensions of the technologies in question are somewhat different.

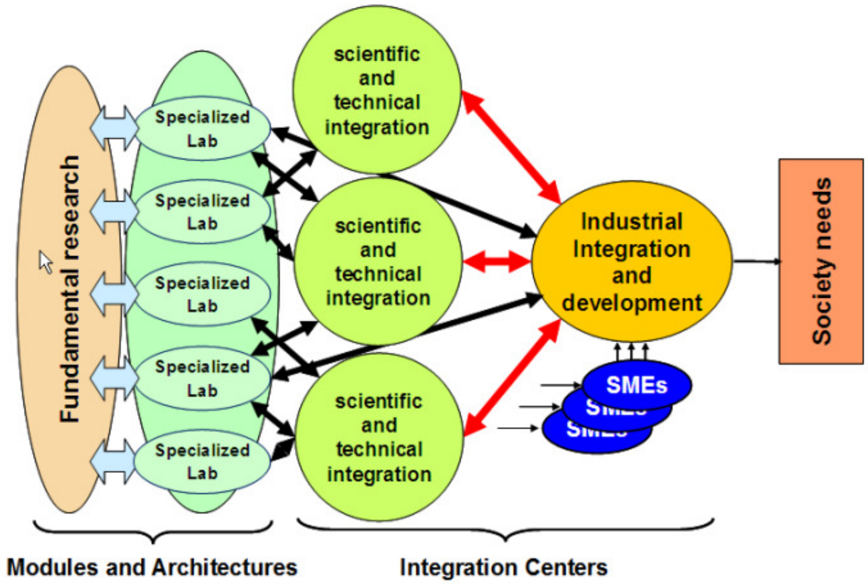

Fig. 7 – Links between Industry and Academia in MNT (image from PRINS documents).

1.6.1 MNT Domain Classification

More specifically, MNT ICT technologies span five major domains:

a) *Micro/Nano-electronics*. This technology enables the fabrication of microprocessors, memories and RF communication devices. The investments for a state-of-the-art facility have become extremely heavy (multi-billion €), so only a few players are left in this field. In Europe, CEA-LETI, IMEC and some Fraunhofers are the only RIs that remain.

b) *MEMS*. This domain originally started by using microelectronics technology for making devices that interact with a non-electrical domain. Examples of MEMS devices are accelerometers, microphones and fibre-optical switches. Many RIs exist in this domain, whilst the number of new applications of the technology is rapidly diminishing.

c) *Electro-optics*. This technology is similar to microelectronics, but uses III-V semiconductors to fabricate solid-state lasers, LEDs and some high-frequency circuits. This technology uses very old micro-electronics scale technologies, combined with complex machines for the growth of the semiconductors.

d) *Top-down nano-technology*. This technology uses some very complex machines – e.g. electron beam direct writing tools (*e-beam*) – to define extremely small nano-scale structures that can interact at the molecular level. One example of an application of such technology is that it can be used to create a filter that consists of a membrane in which very small (nm) pores have been defined using an e-beam.

e) *Bottom-up nanotechnology*. This technology also defines structures that interact at the molecular level, but uses combinatorial chemistry to define them. An example is a filter which has an alumina membrane that has been made nanoporous by anodization.

1.6.2 MNT RI Collaborations

The *Micro and Nano Technology Europe* (MNTEurope) project has led to the creation of a unique distributed platform for research and development in the field of Micro and Nano Technologies (MNT), based on the national facilities of its five partners, Léti (France), CSEM (Switzerland), IZM-M and IISB Fraunhofer (Germany), IMEC (Belgium) and Tyndall (Ireland) research institutes; the *MNT Europe Alliance* currently offers a list of 227 MNT *technology capabilities* available in their respective national facilities.

The *European Integrated Activity of excellence and Networking for Nano and micro-electronics Analysis* (ANNA) offers analytical support to Research Institutions and also production facilities for SMEs and Industries. The ANNA consortium comprises twelve partners from seven European countries. The strength of this network is based on the combination of complementary backgrounds of the partners. The consortium covers research institutions, universities and industry. In the ANNA model, access to the (distributed) RI MNT infrastructure is granted following a selection of the proposals by "peer review". Costs for the access during the project were covered by the EC.

The ANNA MNT Research Infrastructure, available from 2008 to 2010, had the following characteristics:

- more than *one thousand days* of EC-funded *Transnational Access* to *eighteen facilities* (laboratories, clean room, metrology) was offered at eight locations,
- access to ANNA instrumentation and analytical services was either in person or remotely by suitable electronic communications,
- *users of the infrastructure* were researchers from small & medium enterprises, large scale industry, research centres, or universities,
- interested users apply for research access by submitting a short *project proposal*,
- selection of user proposals is by "peer review" on the basis of *the scientific/technical merit*. Priority was given to first-time users and to users in countries without a similar infrastructure.

The *Pan-European Research Infrastructure for Nano-Structures* (PRINS) project aimed at enabling European innovative research for the ultimate scaling of electronics component and circuits. The PRINS platform was searching for the convergence of "top-down" microelectronics technology with "bottom-up" methods derived from fundamental disciplines such as materials science, physics, chemistry, biotechnology and particle electronics. This research infrastructure aimed at a cross-disciplinary fertilization of academic and industrial competences in the areas of nanoelectronics, nanosystems, nanobiology, nanophotonics, etc. The PRINS consortium consisted of three leading research centers (IMEC, CEA-LETI and FhG-VμE), four industrial partners (ASM-L, Infineon, NXP and STMicroelectronics) and the Public Authority/Funding Agency of the Region of Flanders (EWI).

According to the experiences of the three projects, future MNT RIs should be distributed and created as an extension of current Research Organization facilities, as it appears improbable that, in the short to medium term a strong, public-supported, effort to create a centralized MNT RI is feasible. The collaboration efforts were mainly directed towards the definition of common agreement for access to the processing equipment and to consolidate a common view of all offered capabilities. Intellectual Property (IP) and legal aspects are of importance in this sector and must be clarified via contract agreements and definite procedures.

We report in the following the outcomes emerging from the PRINS Final Report:

R1. *Governance Structures*. It was thought that it would be advantageous to have for the implementation phase of PRINS, a *Scientific Advisory Committee* (SAC), *an Industrial Advisory Committee* (IAC) and a *Public Authorities Advisory Committee* (PAC) to support the management of the RI. During the execution of the PRINS project thought was given to the composition of the different committees and it was noticed that while there is strong interest from both the scientific and the industrial sides, the Public Authorities were *quite reluctant* to contribute to the preparatory phase but rather interested in discussions based on a complete concept.

R2. *Intellectual Property (IP) and Legal Aspects*. Depending on the *access mode* used during the operation of the RI, the IP ruling is different. An important issue is whether the scientific interactions are only based on the use of *background information* or whether *new information* is generated during the access period by the visiting and/or hosting organization or jointly by both. In the first case the IP remains within the hosting institution, while for the latter *IP sharing models* are more appropriate. The ROs involved in PRINS each had a different operational modes generic document for covering the IP ruling. These documents had to be fine tuned on a case by case basis in order to ensure optimal protection of the IP of the different partners involved.

R3. *Processing Equipments & Technologies*. The *interest* of the scientific community to use the access/hosting offered by MNT RIs will strongly depend on the type of processing equipments made available. Each of the MNT RI access nodes should have detailed procedures in place related to the practical aspects of giving access, covering also issues such as contamination control and compatibility with standard operation of the facility.

R4. *Existing MNT RIs / Benchmarking*. A *benchmarking exercise* has been performed in order to compare the PRINS RI with existing similar initiatives worldwide. This exercise included a detailed study of three mature initiatives like SEMATECH (US), NNIN (US), and NDL (Taiwan), the Taiwanese Industrial Technology Research Institute (ITRI) and the Nanotechnology Initiative in Korea. The different initiatives gave a successful answer to a specific purpose in their particular ecosystems but *failed to serve as overall prototypes for the PRINS infrastructure initiative*. SEMATECH and ITRI are mainly focusing on the semiconductor industry, while NNIN is operating at university level but lacks access to state-of-the-art technologies. The Korean initiative brings major stakeholders together in a long term strategic plan at national level. Therefore, transferability of best practices from these different initiatives *can only be evaluated for specific cases*.

R5. *Scientific Communities*. The *scientific collaboration* can be based on: hosting research teams in the RI and/or building up common R&D laboratories, mobility of researchers, and cooperation in scientific and technical projects, providing academic research teams with basic materials/data for their own research programs. Different aspects of these operation modes have been considered and worked out. The three RO’s involved in PRINS each already have experiences with the different forms of access/hosting. An important aspect, however, regards the *associated financial implications*. A workshop within the scientific community pointed out that preference is given to *free access*. This could be solved by allocating a dedicated budget coming from the European Commission. The Member States could also at national level allocate a dedicated access/hosting budget.

R6. *Strategic Objectives*. The access given by the MNT RIs should enable the academic community to perform advanced research in all main MNT domains, i.e., 'More Moore', 'More than Moore' and 'Beyond CMOS' and to remain at the leading edge of the research world. MNT RIs should *include the ROs needed* to ensure an optimal and cost-effective use of the research facilities available in Europe. Micro- and nano-electronics research is becoming *multi-disciplinary* and therefore requiring the combination of different skills.

R7. *Funding and Legal Structure*. An MNT Research Infrastructure has some special features and requirements. Research in nanoelectronics is to a large extent *application-oriented* and is performed *in close cooperation with industry* in order to take into account potential *manufacturing aspects*. A technology-oriented Research Infrastructure needs access to *very expensive* state-of-the-art equipment and technologies, which have to be updated on a time scale of three-five years. Due to the high equipments costs and high operating costs of the clean room facilities, such an infrastructure would be extremely expensive. Building a new centralized RI facility would cost *several billion euros* and would therefore require a *strong and long term financial commitment* of the funding agencies and/or Public Authorities of the Member States and/or Associated States. It appears not to be feasible for an agreement between Member States to agree on a new *centralized* RI for nanoelectronics – on the other hand, a *distributed* approach based on combining the existing facilities at different locations is a realistic starting base to create a nanoelectronics RI. Presently, for each of the three ROs in PRINS the financial model only to a small extent relies on the support of the Public Authorities and mainly relies on contract research with industrial partners on a world-wide scale. Basically the funding is based on *public-private-partnerships* and *contract research for the industry*. Keeping in mind the industry oriented business model and restricted contribution and interest of the Public Authorities and funding agencies, it was concluded that forming a separate legal entity, e.g. based on the ERIN concept, *would not be feasible and also not desirable*. All the intended activities can easily be executed by a MNT RI under which the access nodes *keep their independent legal status* and collaboration agreements are signed between the ROs.

MNT research infrastructures have several points of contact and share several characteristics with *nanobiotechnology*, so it could be useful to learn from parallel nanobio experiences. The *EuroNanoBio study* has shown that a distributed infrastructure based on a coordinated pan-European alliance of research driven clusters is the most efficient instrument for fast and efficient translation of nanotechnological discoveries into biological and medical applications. We highlight in the following (with italics) the relevant facts emerging from the fifteen recommendations reported in the study:

R1. To cover the large range of scientific disciplines involved in nanobiotechnology and the diversity of application areas, a European infrastructure has to be built on *regional* *nanobio clusters*, which have *world-class facilities and expertise* with high levels of engagement between industry and academia.

R2. The nanobio clusters *need to be connected and coordinated to share knowledge and equipment* and to cover the whole value chain in specific application areas of nanobiotechnology such as environment or medicine, for example.

R3. A *dedicated infrastructure management* should improve the engagement between academic disciplines, research centres and industry inside and between the involved clusters.

R4. *Clear technical roadmaps* for each of the application areas within nanobio should be defined *to provide a catalyst for collaboration between industry and academia* within the infrastructure and where new exploitation pathways for groundbreaking science can be defined and supported.

R5. Experts should be encouraged to work collaboratively with science departments, research institutes and industry to help explore *Ethical, Legal and Social Aspects* (ELSA) of developing nanobiotechnology, thereby enabling early decision making on the probability of commercialisation in a socially and ethically responsible manner.

R6. Set-up and upgrading of clusters will require *local, national and European political support and funding* which would be supplemented by *private investments at a subsequent more mature stage*.

R7. A *European reference centre* is needed for *characterization and toxicology studies of nano-objects*, which can be accessed by all nano-object producers and users from academia and companies similar to the Nanotechnology Characterisation Lab at NCI/USA.

R8. A *European centre for Risk and Safety Management* should be established, which provides information and advice about the handling of nano-objects and protection measures to SMEs and universities that cannot afford expensive risk assessment.

R9. *Clusters should especially help and support SMEs* to articulate their needs and interests to regulatory and standardisation bodies.

R10. The infrastructure *should provide pools of experts* and *professional communication tools* necessary for engagement with the public.

R11. Promotion of the capabilities of nanobiotechnology to SMEs and clinicians should be facilitated by *showcasing examples of successful exploitation of nanobiotechnology*.

R12. Engagement of the European infrastructure with nanobio clusters and research centres *outside Europe should be encouraged*.

R13. The *highly interdisciplinary nature of nanobiotechnology* requires the integration of dedicated nanobio modules *preferably at the MSc or PhD level*.

R14. Because nanobiotechnology touches on many important wider issues, *teaching an understanding of ethical and social aspects* together with *training in science communication and public engagement* should be *included* at the MSc and PhD level.

R15. Due to the rapid development in nanobiotechnology *targeted education and training programmes for in-career training need to be developed*.

1.7 The Research Data Infrastructure Framework

Data infrastructures are on hand a very new part of e-Infrastructures open to many activities but on the other hand, although many data infrastructures exist, they are not always coordinated at national levelor at European level.

Two most important recent events in the domain of European data infrastructures are the report from the High Level Expert Group on Scientific Data: "Riding the wave: How Europe can gain from the rising tide of scientific data" and the recently commenced EUDAT project.

For completeness the "older" activities in the domain of data infrastructures are described.

The report describes the large amount of information that will become available between now and 2030, makes a wish list of how an ideal data infrastructure will look, reflects on the benefits of data infrastructures for researchers and citizens and also tackles challenges that need to be overcome. The report includes a series of recommendations to make data infra structures with a European view in mind and addresses these recommendations to national governments, the European Commission and researchers alike.

EUDAT is a project to create a *collaborative scientific data infrastructure* with the capacity and capability for meeting future researchers' needs in a flexible and sustainable way, across geographical and disciplinary boundaries. EUDAT will undertake the first comprehensive European review of the approaches to deployment and use of a common and *persistent* data e-Infrastructure, the services being built and delivered on top of this infrastructure and the limitations of its services. EUDAT will be ensuring the *authenticity*, *integrity*, *retention* and *preservation* of data deposited by its users, especially those marked for long term archiving.

The EUDAT consortium brings together e-infrastructure providers, research infrastructure operators, specialist application/solution providers and researchers from a range of scientific disciplines under several of the main ESFRI themes. The depth and breadth of coverage allows a collaboration between for example, data centres such as JUELICH, KIT, CSC, BSC, SIGMA, SARA and SNIC, that support multiple customer groups as well as three infrastructure-providers (CSC, PSNC and SIGMA) operating 'end-to-end' across the networking, grid, HPC and data-management layers of the e-Infrastructure. Equally important, given the goal of self-sustaining operation within three years, is the involvement of two agencies with national funding responsibilities (SNIC, STFC) able to litmus test the proposition as it emerges. EUDAT is the first project that aims at defining a data infrastructure across domains and disciplines.

As reported by the EC (see e-Infrastructure: "Scientific Data, Accessing Knowledge" leaflet):

"Scientific data have come to the forefront of modern science; it is the new way to express knowledge. Many disciplines are developing into highly data-intensive areas of science such as high-energy physics, astronomy, bio-informatics, genomics and medical imaging. Scientific experiments, observations, theories, models and simulations generate unprecedented volumes of data which will quickly reach the Exabyte scale. This data is stored in complex databases consisting of numbers, text, images, diagrams and formulas."

"The European Commission has been actively promoting and funding digital libraries policies, initiatives and research over the last decade. The i2010 Digital Libraries initiative has been integrated and continued in the new policy framework "Digital Agenda for

Europe", one of the seven flagship initiatives of the Europe 2020 strategy for economic growth. In November 2007 Europe's cultural institutions created a foundation that provides for the organisational underpinning of the service. A demonstration site of Europeana was presented in February 2008. A full prototype was launched in November 2008."

We recall here (the italics are ours) the conclusions of the *Data Management Task Force* (DMTF) reported in the e-IRG Report on Data Management [e-IRG DMTF, 2009]:

- A very large number of *individual* and *domain-specific* data initiatives and scientific databases exist in Europe.
- *Only a small number* of the data initiatives are *federated*.
- Interoperability is de facto limited to a particular scientific domain which applied standard formats at an early stage of the data initiative.
- *Long-term sustainability is a major issue for all data initiatives*, not only is the problem limited to the underlying hardware infrastructure *but also to software* accessing and exploiting the data.
- *Open access* to data has *not yet become a reality* in *all* scientific domains. It is often hampered technically by the absence of search engines, and institutionally by the absence of *clear access policy guidelines*.
- Organisations are moving away from a centralised stand-alone model towards *distributed networks of federated data repositories*.
- New requirements for *cross-disciplinary* research will require interoperability between different disciplines and different types of data.
- *Metadata* is recognised as being paramount for long-term data access and usability (including documenting the research process, not just the data itself).
- The suite of data analysis tools is growing and becoming more complex (i.e. the use of GIS in many fields).
- The focus is on curating data for reuse, not necessarily only for *long-term preservation*.
- *Open access* to data is becoming more common.
- New projects, many of them financed by the EU, will profoundly change the current data landscape in Europe and set standards for the rest of the world.
- Additional efforts should be put into gathering more detailed information about project-specific needs on data management and what kind of activities are required.
- *Communication and cooperation between the data initiatives/projects should be stimulated to achieve better interoperability and reuse of solutions and infrastructures*.

In the Final Report [e-SciDR, 2008] twelve mutually reinforcing recommendation sets for policy measures were summarized towards driving a European e-infrastructure for and of e-Science digital repositories.

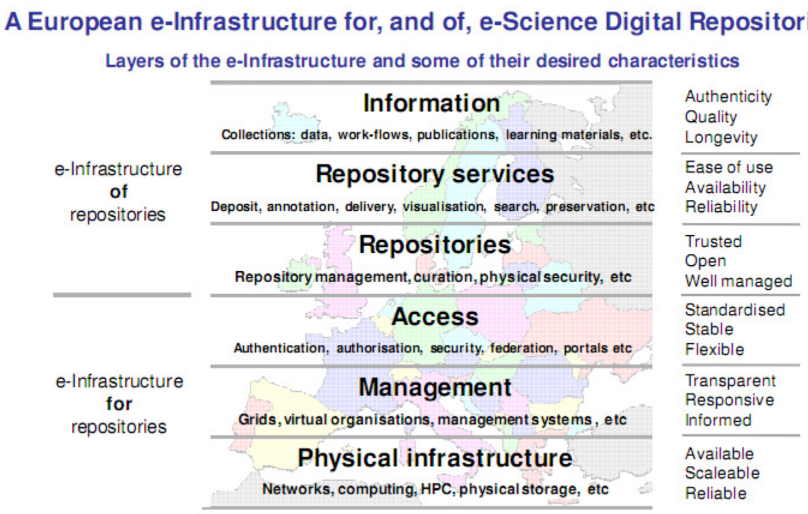

Fig. 8 – The Vision of an European e-Infrastructure for/of e-Science Digital Repositories.

The recommendations were grouped according to perspective and level:

A. *Building an e-infrastructure* for research continuity requires: funding reform, European e-Infrastructure of and for European e-Science digital repositories, governance and management;

B. Engaging *users* and *service providers*: support for data producers, trust and recognition, training and awareness is required;

C. *Providing access* to researchers, educators, students, unaffiliated stakeholders and interest groups: discovery and navigation, *open access* to publicly funded data is needed and legal issues, both local and international, need to be identified and resolved.

D. *Maintaining and preserving information*: collections management, selection and appraisal for sustainability, preservation of digital information are prerequisites.

1.8 The Remote Instruments Access Model

As noted in [ESFRI, 2010], European RIs' *instruments and facilities* – such as *synchrotrons*, databases, *telescopes*, *sensor networks*, and *biomedical facilities* – are an unrivalled asset. In Europe there are *more than 500* such facilities, of which at least 300 have strong international visibility, attracting world class researchers. Again, in [ESFRI, 2010], it is noted that "These RIs represent an aggregate European investment of more than €100 billion. Some 50,000 researchers a year use them to produce 3,000 to 6,000 high-impact research papers – as well as a chain of patents, spin-off companies and industrial contracts." And, "RIs provide the means and impetus to develop a truly sustainable e-infrastructure to store, share and protect digital data. This permits Europe to lead the development of e-science.” Also: “Advanced computer and communications technologies are changing not just the tools of science, but also the methods. *Scientific e-infrastructure permits researchers in one place to undertake experiments on RIs remotely, in real time*; to model, simulate or infer conclusions from vast data sets; and to collaborate with researchers of widely different backgrounds and disciplines. Some[5] see this as creating a “fourth paradigm” of science – beyond observation, theory and simulation, and into a new realm of correlation to mine new insights from vast, diverse data sets.“

The increasing availability of instruments and facilities (*physical* infrastructures) for researchers, combined with the flexibility and versatility of the deployed e-infrastructures (*digital* infrastructures) is opening a completely new paradigm for the undertaking of science experiments, where scientists will have the possibility of interacting remotely with their instruments, reducing the geographical distances thanks to the *Remote Instruments access model*.

The EU FP7 *Deployment Of Remote Instrumentation Infrastructure* (DORII) Project[6] complements the Grid middleware with the *Instrument Element* (IE) in order to enable remote access to instrument resources such as:

- seismic sensors and actuators for *Earthquake observation*,
- conductivity, temperature, depth (CTD), pressure, oxygen and turbidity *sensors*, digital cameras, *Autonomous Underwater Vehicle* (AUV) and others for *Environmental sciences*,
- X-ray scattering, detectors, beamline, synchrotron radiation in medical application, for *Experimental science*.

IT technology, provided by partners who recognised the initial requirements of scientific communities, have worked to deliver the functionality developed in projects which had a focus on accessing remote instrumentation (GRIDCC, RINGrid), on interactivity (int.eu.grid), on software frameworks for application developers (g-Eclipse) and advanced networking technologies (GN2) with EGEE based middleware.

[5] See: [Hey, Tansley, Tolle, 2009].
[6] See: www.dorii.eu.

1.9 The Future Internet Service-Oriented RI Vision

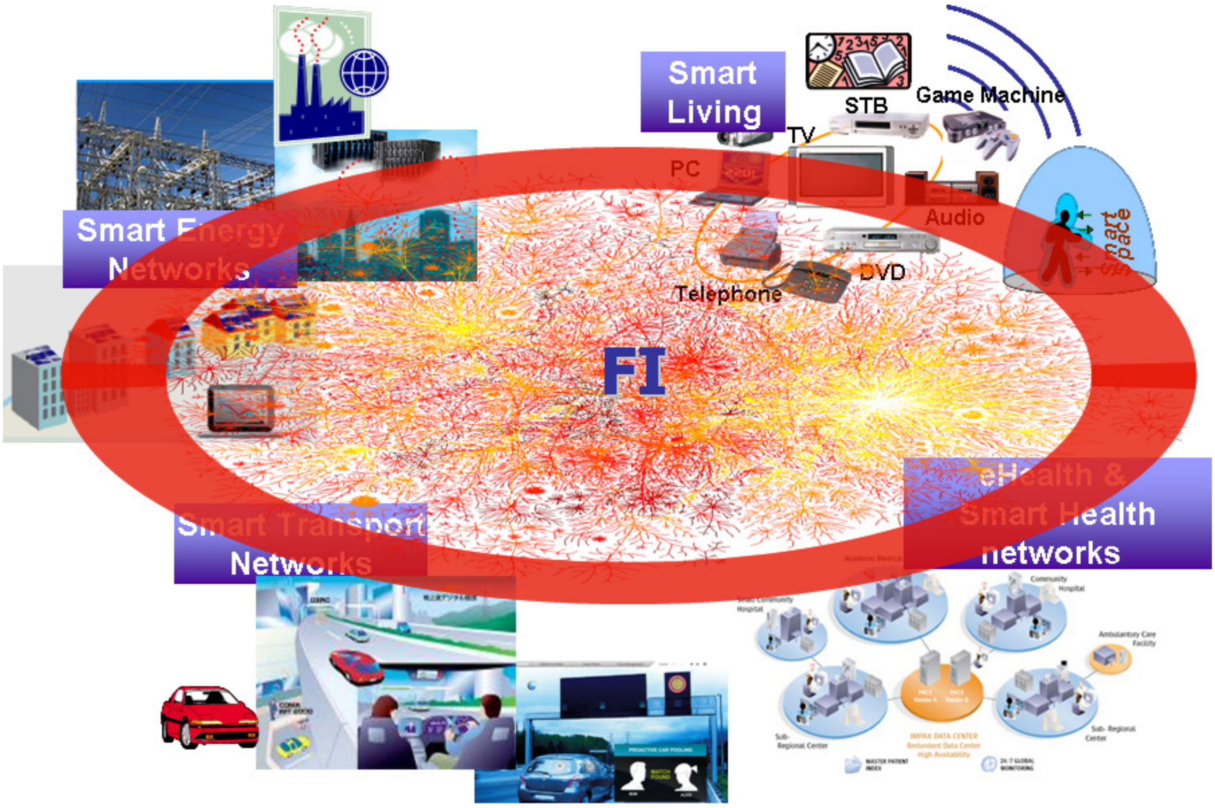

Fig. 9 – The Future Internet Vision.

The following *usage trends* for the *Future Internet* (FI) were suggested in the [WGFI, 2008] report:

1. *Mobility*, needs a *native support* in an infrastructure conceived for fixed usages;
2. End to End *very high rate throughput*, reaching the limits of current Internet protocols, not designed for ultra broadband scenarios;
3. *Security and Trust, Privacy*, to be supported in FI directly in the service and network infrastructures (*privacy by design*);
4. The "internet of things" – device connectivity, coupling of virtual world data with physical world information (RFID, sensors), needing an FI with high network architecture scalability and new protocols and service architectures to support device generated traffic and service discovery;
5. *User generated services*, as a follow up to user generated content, via new FI service architectures enabling dynamic, secure and trusted service compositions and mash ups;
6. *3D* becoming mainstream, requiring a FI capable of resource intensive usage of computing, networking platform and new standards;

7. *Negotiated management and control of resources*, negotiated SLA's, to be enabled via FI dynamic and predictive network management, infrastructure observability and controllability;
8. *User controlled infrastructure*, requiring FI ad hoc network and service composition.

New, FI-related infrastructures, still to be defined, should also satisfy these requirements.

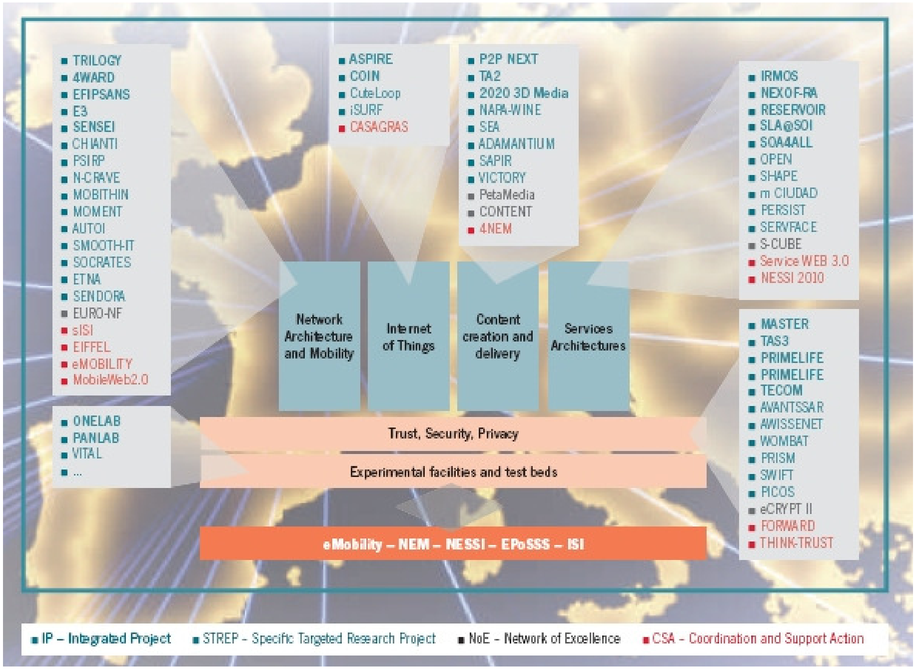

Fig. 10 – EC-funded Future Internet projects (image from www.future-internet.eu).

With a view towards future possible FI Research Infrastructures, the same document discusses the main issues that need to be addressed for an EU wide approach to FI:

- *Networking*. Networking of *actors*, within and across the national boundaries, as in Eureka and Framework Programme; networking of *knowledge*, for a multi-disciplinary approach, such as in the "Living Lab" and open innovation models; networking and *development of skills*, considering the experience of the *Knowledge and Innovation Communities* (KICs) of the *European Institute for innovation and Technology* (EIC); the networking of *users*, who should be involved in early stages of the research in a bottom-up process; networking with researchers outside Europe in a global approach;
- Changing the main target from *technology* to *applications*. The use of experimental testbeds and facilities is suggested as an effective mechanism to take application requirements into account within technological developments. Tighter links between FI and GÉANT (and FEDERICA) could help to federate interconnection of test-beds and sustain a virtuous circuit of innovation for the development and testing of new technologies and solutions. From the deployment perspective, focus should be on applications of public interest whose implementation can be supported by public

administrations. *Health care*, *education*, *e-science*, *e-administration* are typical application areas to consider;

- Addressing *the right mix of actors*. Smaller users, SME's and innovative actors are called upon to actively contribute to the process where significant industry driven work is currently on-going to define a cross domain perspective of the FI;
- The identified *common topics*. Rethinking the *network management* to fit the widest possible application requirements, *mobility* and end-to-end very high speed rates, the "Internet of Things", the "Internet of Services", new models of *massive 3D usages* for entertainment and public applications such as health, information networking (saving, sharing and delivering), security/trust/identity, advancing federated test-beds experimental infrastructure.

2. ICT RIs' Challenges

We recall in the following the main results of [OSIRIS D2.1, 2010].

Research Infrastructure *is a pillar in developing the European Research Area* (ERA). This holds true particularly in the field of ICT Infrastructures. Progress in ICT and its Infrastructures *is a "condicio sine qua non" for most of the other installations*.

ICT Research Infrastructures are identified as *e-Infrastructures* or *Infrastructural Competence Centres* dealing with *Research and Development in the ICT area* and/or *ICT facilities for Research and Development.*

ICT RIs challenges were grouped into the categories:

- *Policies*: public policy, science policy, innovation policy, cooperation policy, technological policy;
- *Resources*: financial resources, human resources, users;
- *Frameworks*: legal framework, organizational forms of e-Infrastructures, intellectual property rights, infrastructure security and data privacy;
- *Scale*: regional dimension and global dimension.

The following challenges were identified:

1) *Infrastructural Competence Centres* (ICCs): creation of European Research Infrastructures or a network of them for the development of *Micro and Nano Technologies* (MNT) *is still a challenge* because cross contamination of samples during exchanges could represent a significant risk to the basic functionality of each contributing institute and because of sample size and processing constraints between facilities. The cooperation between the MNT ICCs could be expected to develop on the basis of an interconnection of complementary and not overlapping capabilities.
2) *Data RIs*: the processing of large amounts of data is rapidly becoming the *most important and challenging ICT problem* when compared, for example, with providing large computing facilities.
3) *Public and science policy*: the wide usage of Research Infrastructures based on the principles of the Fifth freedom is still a challenge for the Member States (MS). The development of Research Infrastructures will challenge changes in the existing science policies of Member States and also their ability to coordinate national policies in the overall context.
4) *Innovation & Cooperation policy*: the organization of cooperation between scientific institutions and business companies in the framework of Research Infrastructures is especially challenging because the interests of these two groups and their attitude to multinational cooperation may be different.
5) *Technological policy*: creating a common technological policy for e-Infrastructures is a challenge for ICT researchers and for the ICT Research Infrastructures.

6) *Standards*: standardization is necessary for the development of e-Infrastructures and *its development is a serious challenge*.

7) *Usability*: the "ease of use" principle should be *a primary objective* as the ICT RIs enlarge their user base from the current, more skilled communities to future, less skilled communities.

8) *Legal framework*: the work on the creation of the legal framework for Research Infrastructure activities has been started but remains a challenge. The law for setting up the *European Research Infrastructure Consortiums* (ERIC) has been adopted but the rules for how it can actually function are still incomplete.

9) *Intellectual property rights*: the deployment of results from Research Infrastructures is challenging even for experts in *Intellectual Property Rights* (IPR).

10) *Infrastructure security*: ensuring Research Infrastructure security (i.e. solving security issues related to the RI and the data/results it handles and disseminates) is challenging, and taking this into account should be regarded as an important part of any RI project.

11) *Data privacy*: any projected Research Infrastructure should be audited regarding *personal data protection* risks.

12) *Regional dimension*: the partnership role of *regional level specialised Research Infrastructures*, if *coherently integrated* into the whole ERA, would include better exploitation of results, training young researchers and a broad promotion of research performed at the large facilities.

13) *Global dimension*: ensuring the long-term commitment and continuous financial support required for those e-Infrastructures operating on a global scale is difficult to achieve and will be challenging from individual government policy and their financial perspectives in cross border scientific support arrangements

14) *RIs strategy*: the development of Research Infrastructures will continue, but organizational changes may be expected.

2.1 General characterization of the PA-RIs cooperation

For general insight into problems concerning the cooperation of Public Authorities and Research Infrastructures the results of the OSIRIS Survey may be used. Amongst other issues, this Survey studied the general approach taken by the main stakeholders to cooperate with ICT Research Infrastructures and also identified the main characteristics of the cooperation. The specific problems of infrastructures, of course, were not investigated. The results of the Survey are overviewed in the deliverable D2.5 presented in two parts, and this discussion is based on these results.

The selection was made in order to have a representative sample of the European bodies that manage the Research Infrastructures – not as individuals. The (hence small) sample of respondents to the OSIRIS Survey also comprised representatives from Public Authorities. The respondents were divided into several groups, and the Policy group consisted of respondents who either belonged to public authorities with competencies in the ICT sector and either had experience in policy, running, planning and/or financing ICT RIs or were engaged in cooperation with existing ICT Research Infrastructures in other ways. Their opinion concerning cooperation may reflect to some extent the approach of European Public Authorities. At the same time the opinions of the other respondent groups – Operation and Users – may present some understanding of the perception of the PA activities in these groups.

2.1.1 Priorities of Research Infrastructures

The respondents were asked to indicate the main areas which they represent: they were able to select several Infrastructures from six proposed types. On average respondents checked about two RIs as being most relevant to their activities. The vast majority of the respondents regarded themselves as being closely connected to *Computing Infrastructures* (63.6%) or to *Communication Infrastructures* (52.7%). The Policy group showed similar results, as is seen from Table 1, presenting the percentage of RIs types checked by the respondents of this group. The last column presents the appropriate percentage for all respondents.

Clearly, the Policy group considered the Computing Infrastructures as being the most important Infrastructures – the respondents of this group checked these Infrastructures in 73.10% cases. The results on the table would suggest that the group focussed on the care of the Computing and Communication Infrastructures (grid, HPC and similar). The Future Internet, Data Infrastructures and Instrumental Infrastructures appeared to be somewhat less important to the Policy group, and MNT Research Infrastructures were even less highly regarded.

Table 1 – Percentage of checked RIs by Policy group.

Infrastructures	Policy group	All resp.
Computing Infrastructures	73.1%	63.6%
Communication Infrastructures	53.8%	52.7%
Future Internet	26.9%	25.5%
Data Infrastructures	26.9%	18.2%
Instrumental Infrastructures	19.2%	14.5%
Micro and Nano Technology RIs	15.4%	18.2%

It is apparent that such a distribution of priorities does not correspond to the effective relevance of different RIs for science communities, but rather it is more related to the RIs' involvement of government agencies.

Additionally, the Survey also studied the usage of ICT Research Infrastructures by respondents. The survey results indicated that in total more than half of them used the RIs in their research. The usage of these RIs by the Policy group was less active – 39.3% of them used these infrastructures, 35.7% did not and to 25.0% the usage was not applicable. It suggests however, that, generally speaking, Public Authorities also actually use the Research Infrastructures in a practical sense and are indeed engaged in the process of development and collaboration.

2.1.2 *Assessment of scientific cooperation*

The Survey also studied the opinion of respondents regarding the level of scientific cooperation in their sector of the ICT field in their countries. These results can be used to evaluate the general attitude towards cooperation in-so-far as it affects the development of Research Infrastructures.

The respondents could choose one out of five possibilities in order to characterize their assessment of scientific cooperation in the ICT sector of their respective countries. The chosen answers of the Policy group are shown in Table 2.

It may be concluded that the opinions of the Policy group differ somewhat from the averaged opinion of all respondents. Whilst most of the respondents thought that the cooperation which they experience is satisfactory, the opinions of the Policy group are somewhat extreme: compared with the perceptions of all respondents, more respondents from this group perceived cooperation to be good in the whole ICT sector including RIs and simultaneously more of them evaluated it as being poor. The high percentage of the opinion about poor cooperation in the Policy group may be considered as a form of implicit self-criticism, as in reality this group could do a lot towards improving scientific cooperation.

It should be mentioned that the respondents in their comments to the answers identified two main obstacles for the development of scientific cooperation in the ICT sector. The first is the fragmentation of organisations on the national level i.e. a rather large number of small organisations and enterprises with fewer resources for research and also on the European level, where the national research (data, MNT, …) infrastructures should better cooperate, following the example of Géant and EGI. It was proposed to promote networking actions at national level and to coordinate them with European measures. It is very important for the cooperation to define national research programs and national roadmaps for RIs, taking into account the explicitly expressed needs of users.

Table 2 – Assessment of cooperation by Policy Group.

Assessment of cooperation	**Policy group**	**Total**
Satisfactory	28.6%	42.1%
Good in the whole ICT sector	39.3%	29.8%
Good only in the Research Infrastructures	10.7%	12.3%
Poor	17.9%	10.5%
Good in the ICT sector, except Research Infrastructures	0.0%	1.8%[7]
No opinion	3.6%	3.5%
Total	100.0%	100.0%

[7] The cooperation in ICT Research Infrastructures is considered better than in the whole ICT sector by virtually all respondents. The small percentage with a different opinion – 1,8% – corresponds to just one response in our sample, belonging to the Usage Group and coming from a country outside Europe.

As the second obstacle to research cooperation, the competition between institutes within countries for limited funding was mentioned. It should be added, however, that the EU science policy in the Framework programmes was always to support collaborating research institutions of several countries working on sufficiently large projects, and that this stimulated the research cooperation on the EU level.

2.1.3 Coordination of priority research themes

One of the ways in which scientific cooperation can be developed is by the clear definition of priority research themes in the country. If the priority research themes are set up without coordination with other countries it is difficult to expect good cooperation between those countries. The Survey asked respondents if the priority research themes in the ICT sector of their respective countries are coordinated with those of other Member States.

The appropriate results for the Policy group are presented in Table 3.

It was rather surprising that the Policy Group respondents had a rather low awareness about the existence and requirement for the European coordination of the priority ICT Research themes. The percentage of *No* answers to this question amongst groups was highest in this group as was the percentage of respondents who had no opinion. Only one in four respondents admitted to an awareness of the existence of this type of coordination. The same indicator was about two times larger for this group than for other groups.

Table 3 – Percentage for coordination of priority research themes.

Coordination	**Policy group**	**Total**
Yes	25.0%	38.6%
No, they are set by the science policy of the country	46.4%	38.6%
No opinion	28.6%	22.8%
Total	100.0%	100.0%

The comments of the respondents in free form allowed the conclusion that coordination mechanisms are widespread and well developed, and mainly based on participation in the Framework Programmes, eIRG or ESFRI and in various European RI projects like GÉANT, PRACE, EGI, Artemis and others.

2.1.4 Influence of RI on research themes

The influence of existing or planned RIs on the priority research themes in a country can be considered as both a manifestation and result of the research cooperation relevant to the RIs. Therefore the respondents were asked to assess the observed influence.

The results are presented in the following table.

Table 4 – Percentage of influence of RIs on research themes by main activity groups.

Do RIs influence research themes?	Policy group	Total
Yes	50.0%	52.6%
No	28.6%	21.1%
No opinion	21.4%	26.3%
Total	100.0%	100.0%

Generally speaking, this question surprised respondents and many of them had no definite answer. The opinions of the Policy group were approximately the same as that of the typical answers of all respondents. However, the 50.0% of the positive answers can not be considered as an indicator of great influence of RIs on the development of research themes.

The answers to a previous question concerning the coordination of priority research themes in the respondent countries with other MS were used to establish whether the coordination of research themes increased the influence of RIs on the research themes or had no effect. The results show with certainty that the research themes that are coordinated with other Member States are influenced by the existing or planned Research Infrastructures. 72.7% of the respondents who agreed that the priority ICT research themes are coordinated with other MS found that RIs influenced the priority research themes and only 13.6% thought that such influence did not exist.

In the comments the respondents mentioned the ways in which this influence is implemented. They reported about the usefulness of the ESFRI roadmap, various EU projects such as FIRE, PRACE, EGI, GÉANT, etc.

At the same time, it was also mentioned that there is interplay between investment into research infrastructures and relative stability of related research teams and laboratories because good and widely orientated projects using European range RIs may be very supportive for the survival of research institutions.

2.1.5 Development of EU ICT Research Infrastructures

The respondents also assessed the level of country contributions to the development of the EU ICT Research Infrastructures. This assessment can also serve as an indicator of cooperation activities. The options available were – wide participation, participation by some institutions only, or no participation. It was proposed also to identify opinion on the contribution of Member States to the development of each type of Research Infrastructures. The results for all respondents are overviewed in the following table:

Table 5 – Evaluation of the country contribution in development of RIs.

Evaluation	*Computing Infrastructures*	*Communication Infrastructures*	*Data Infrastructures*	*Instrumental Infrastructures*	*MNT Research Infrastructures*
Participate widely	28.1%	49.1%	19.3%	12.3%	8.8%
Only some institutions	54.4%	28.1%	43.9%	26.3%	40.4%
No participation	3.5%	1.8%	8.8%	15.8%	7.0%
No opinion	12.3%	17.5%	24.6%	42.1%	40.4%
No answer	1.8%	3.5%	3.5%	3.5%	3.5%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

This table can be compared with the following table that presents the analogous percentage values for the responses by the Policy group only:

Table 6 – Evaluation of the country contribution in development of RIs by Policy group.

Evaluation by Policy group	**Computing Infrastructures**	**Communication Infrastructures**	**Data Infrastructures**	**Instrumental Infrastructures**	**MNT Research Infrastructures**
Participate widely	39.3%	50.0%	25.0%	17.9%	10.7%
Only some institutions	53.6%	28.6%	60.7%	35.7%	42.9%
No participation	0.0%	0.0%	3.6%	7.1%	7.1%
No opinion	7.1%	17.9%	10.7%	39.3%	35.7%
No answer	0.0%	3.6%	0.0%	0.0%	3.6%
Total	100.0%	100.0%	100.0%	100.0%	100.0%

An answer of "wide participation" was highest for the Communication Infrastructures and second highest for Computing Infrastructures both for the responses of all respondents as well as for the Policy group alone. At the same time, the response to "wide participation" and the "participation only by some institutions" was highest for the Computing Infrastructures, again for both the total values and Policy group.

It is seen also that the Policy group has a higher opinion about participation in all types of Research Infrastructures especially in the development of Data Infrastructures, though the level of participation was reported as rather low especially for the Instrumental Infrastructures and MNT Research Infrastructures. Generally the values of the level of participation in the development of various RIs correlate with the importance range of RIs mentioned above.

2.2 Relevant Survey Results

We report here the main results from the survey submission and subsequent analysis:

R1. The *level of scientific cooperation* in the ICT field in the EU countries is *satisfactory*. It is better in the Operation group than in the Policy group.[8] The development of the cooperation is hindered by fragmentation of the sector and by research competition.

R2. The *priority research themes* in the ICT sector in countries with other Member States are coordinated insufficiently, the level is especially low in the Policy group, though coordination mechanisms exist and are used.

R3. The existing Research Infrastructures have no serious influence on the organisation of research activities, especially in the Policy group. However, the research plans of some countries have identified the need for new RIs, especially in the EU12 group[9]. The Policy group in general does not feel that this is necessary.

R4. The European coordination undertakings, advisory bodies, working groups and organizations from the ICT Research Infrastructure point of view *are considered to be important* though not all their activities are considered necessary.

R5. *Paid research services* could be provided by ICT Research Infrastructures mainly in cases where no other funding sources exist. Generally the respondents think that only open access based on peer review process guarantees the best and most efficient use of RIs.

R6. To *improve the cooperation between ICT RIs and industry*, it is proposed that PR activities, visibility and information, introduction of appropriate tax incentives, improvement of the legal environment, establishment of the full innovation chain within the value creation process, participation in common projects and the making of all non-exclusive licenses available royalty free across the entire EU should be intensified.

R7. Using *intellectual investments* as contributions to the development of the ICT Research Infrastructures is supported by approximately half of the respondents; the other respondents have some doubts about the acceptance of such a model.

[8] In WP2, the *Operation Group* was formed by the persons operating national/regional ICT Research Infrastructures and the *Policy Group* was mainly represented by those who have experience in policy, running, planning and/or financing ICT RIs and Public Authority with competences in the ICT sector; the *Usage Group* consisted of ICT RI users.

[9] Country groups were defined in WP2 as: *EU15*, the 15 countries that were members of the EU before the enlargement on 1st May 2004; *EU12*, the 12 countries that joined the EU over the past years; *other countries*, the associated countries and other countries.

R8. *Three out of ten* respondents accepted *Public Private Partnership* (PPP) in the creation of the RIs without any reservations. A further four out of ten thought that *the existing legal framework is not sufficient and should be improved*.

R9. Two out of ten respondents thought that the scientists experienced no difficulties in working with the Research Infrastructures, and more than half of the respondents saw the necessity for *additional training* orientated to usage of RIs. It may be concluded that *the additional training can and should be provided by the Research Infrastructures themselves*.

R10. The prevailing management systems are *different* for the various Research Infrastructures. The RIs can be divided into *two groups* – the HPC, Grid Computing and Communication Infrastructures with management being primarily at *national level*, and the *Data, Instrumental and MNT Research Infrastructures* where management occurs on the site/laboratory level. The management by an international body was indicated only in some responses, specifically in the Grid Computing and the Communication Infrastructures.

R11. More than half of respondents accept *open access* to the Research Infrastructures for all researchers. One out of five respondents would prefer open access *to academic researchers only*.

R12. One out of five respondents thinks that the existing legal framework of the *Intellectual Property Rights* (IPR) is sufficient, others hope for serious or minor amendments. The necessity for serious elaboration of the existing legal framework is much more pressing in the EU12 group.

R13. The most supported legal form for the RIs is *ERIC*.

R14. Data security *is not sufficient in all RIs*; the best situation is in the *Computing Infrastructures*. Using a federative network is the most desirable way to improve the situation.

R15. The *existing privacy* in the RIs is at a higher level than security although the enforcement of additional privacy measures is considered necessary.

R16. Half of the respondents participate in regional projects. The EU12 group and the Policy group are more active in regional cooperation. The preferred infrastructures for regional development are *Data Infrastructures*. The EU12 group also supports regional development of the Computing Infrastructures, and the EU15 indicates the MNT Research Infrastructures as being the best candidate for regional development

R17. The majority of respondents, but certainly not all of them, support the creation of an organizational structure for *continuous analysis and recommendations* on the development and maintenance of existing and future European ICT RIs. The best legal form for such an organizational structure could well be ERIC, but Agreement between governments or institutions and Virtual institutes or laboratories could also be considered.

3. Inventory of PA/NC-RIs Collaborations

As a second step, for each RI domain, the project identified specific types of PA/NC-RIs collaborations and briefly described each of them. The result is provided in this section. In order to better investigate the collaborations, we further selected a subset of these for deeper scrutiny. The selected cases were enumerated with a (C#) code, and assigned to individual project partners. For this PA/NC RI collaboration case survey we followed a *common template*:

a) Description of RI
External contributors were asked to report a clear description of the Research Infrastructure.

b) PA/NC collaboration explanation
External contributors were asked to explain the nature of cooperation involvement of Public Authorities and/or "National Champions", aimed at identifying best practice cases and methodologies as the basis for the development of models and suggestions for future RIs.

c) Budget, funding model, economic sustainability
External contributors were asked to report the budget of the PA/NC RI collaboration – both initial and recurrent investments – the running cost, the funding model and economic sustainability of the RI.

d) Governance / Management
External contributors were asked to describe the governance bodies, their relations and the management model.

e) Users and interaction model
External contributors were asked to specify quantitative and qualitative aspects regarding RI users – how many, from where, how they use the facilities and how do they obtain access.

f) Countries and international collaborations
External contributors were asked to specify the countries involved and the international collaboration within and outside Europe.

g) History and evolution
External contributors were asked to outline the RI evolution from the first experiences to the current model, cite relevant previous projects/collaborations.

h) Security
External contributors were asked to specify privacy, protocols, trust, procedures and property rights for the RI.

i) Operations
External contributors were asked to report on the activities needed and the related efforts

The template-based survey was then submitted to (internal or) external contributors. The responses are collected in the Part II document.

3.1 The European and National Network RI environment

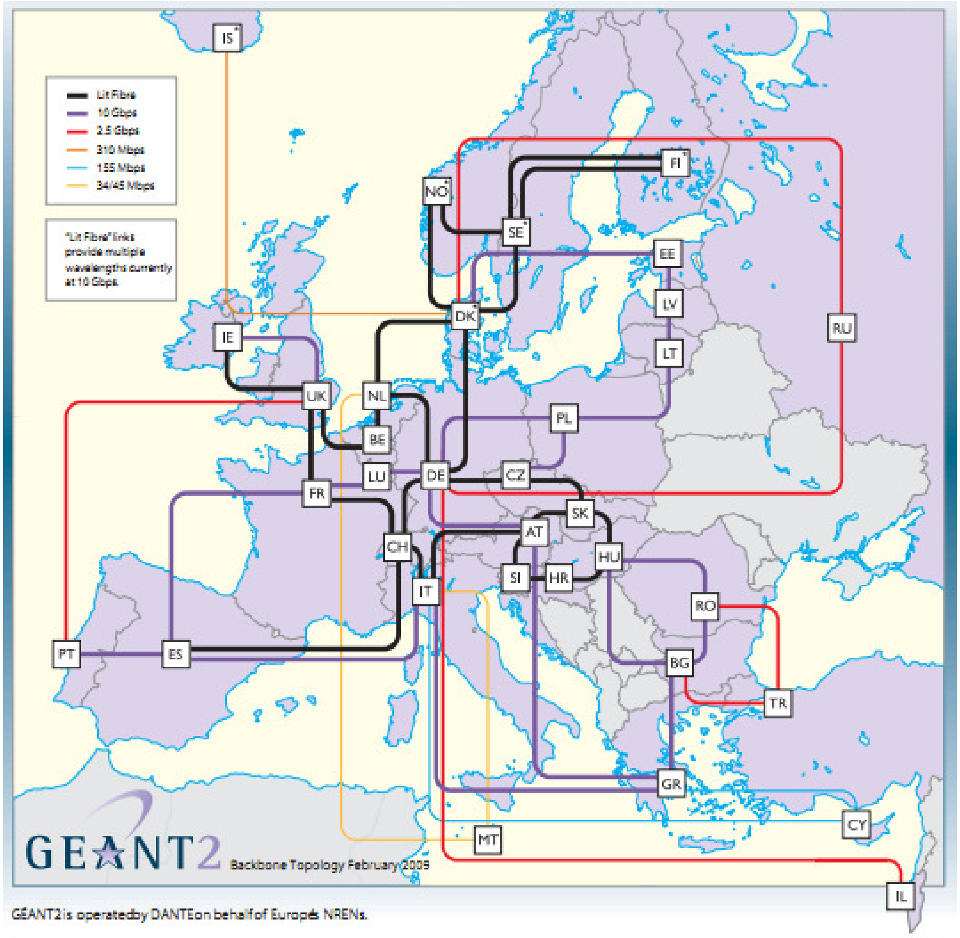

Fig. 11 – The Géant backbone (source: www.geant.net).

3.1.1 National Research & Education Networks (NRENs) Collaboration

A *National Research and Education Network* (NREN) is a specialised network operator and network service providers dedicated to supporting the needs of the research and education communities within a country. It is usually distinguished by support for a high-speed backbone network, often offering dedicated channels for individual research projects and other specialised needs.

Four NREN cases were selected in order to outline the main characteristics of this model of collaboration and detailed in Part II:

- **(C1) SURFnet: Dutch NREN**
- **(C2) Belnet: Belgian NREN**
- **(C3) SUNET: Swedish NREN**
- **(C4) CESNET: Czech NREN**

Key facts on NRENs brought out by the survey

NRENs have been extensively analysed in the 2007 EARNEST foresight study looking at the expected development of research and education networking in Europe over the subsequent five-ten years. The EARNEST work has focused on seven study areas: (1) researchers' requirements, (2) technical issues, (3) campus issues, (4) economic issues, (5) geographic issues, (6) organisation and governance issues, and (7) requirements of users in schools, the healthcare sector and the arts, humanities and social sciences.

We report here the major recommendations of the study:

- there is a *cultural change* in research networking, from providing connectivity to *service provision*;
- provision of services by teams at local and national level should be *more integrated*;
- research and education networking organisations should implement *best management practices*;
- *collaboration* between research and education networking organisations *needs to be intensified*;
- *closer links* should be established with *content providers* and *large user communities*;
- *optical networking* has arrived, but brings new technical challenges;
- *cost-sharing* and *charging* must be fair and *not a disincentive* for innovation;
- *digital and geographic divides* need to be addressed at the political level;
- *user groups with different requirements* need special attention.

The TERENA Compendium is an authoritative reference, published yearly, on the development of research and education networking in Europe and beyond. In the compendium, several statistics can be found on NRENs regarding general, users, network & connectivity services, traffic, high level services, funding & staffing. The 2010 Compendium is the last edition published.

The selected C1-C4 cases are good representatives of the NREN *model of collaboration*, basic and mature instances of the general case of "National Initiatives" to which also NGIs, MNT National Initiatives and HPC National Centres belong.

SURFnet is a non-profit foundation which is owned by the SURF Foundation, the Dutch higher education and research partnership for ICT-driven innovation in which all Dutch universities and research centres collaborate. Belnet is a state service with independent management, managing its own budget and personnel. SUNET is governed by a board appointed by the Swedish research council. CESNET is a not-for-profit association of the Czech public universities and the Academy of Sciences of the Czech Republic.

3.1.2 Network Backbone (GÉANT) Project Collaboration

The GÉANT project is a collaboration between 34 project partners comprising 32 European NRENs, four Associate NRENs, DANTE, and TERENA. Together, GÉANT and the national networks create a common pan-European service area (known as the *GÉANT Service Area*) enabling advanced network services and applications harmonised across GÉANT to be offered by NRENs at local level to institutions, projects and researchers.

NRENs (*National Research and Education Networks*) are specialised network operators dedicated to supporting the data communication needs of the research and education communities within their own country.

DANTE *(Delivery of Advanced Network Technology to Europe*, www.dante.net), established in 1993, plans, builds and operates advanced networks for research and education. It is owned by European NRENs, and works in partnership with them and in cooperation with the European Commission. DANTE provides the data communications infrastructure essential to the development of the global research community and has played a pivotal role in consecutive generations of the pan-European research network. DANTE also facilitates and project manages the creation of inter-NREN structures in regions external to Europe (Asia, South America, southern Mediterranean area).

TERENA (*Trans-European Research and Education Networking Association*, www.terena.org) handles a number of the outreach activities, and supports the co-ordination of the research and development effort amongst project partners. TERENA offers NRENs a forum in which to foster the development of Internet technology, infrastructure and services to be used by the research and education community. It also creates a platform for exchanging research and practical experiences both between NRENS and with hardware solution manufacturers. TERENA also establishes working groups, e.g. EARNEST in the past and now ASPIRE, to study, assess and report on best NREN practice. Its core business is to bring together managers, technical specialists and other people in the research networking community with their counterparts from other countries in Europe (and even further afield), mobilising the expertise and experience of hundreds of professionals in the research and education networking area. TERENA creates a platform for exchanging research and practical experiences both between NRENS and also with hardware solution manufacturers.

We selected the GÉANT collaboration model as a specific case.

(C5) GÉANT Project Collaboration

Key facts on GÉANT brought out by the survey

The GÉANT network and associated programme of activities is funded by the NRENs and the European Commission within the GÉANT Project (GN3) contract, which is part of the EC's Seventh Research and Development Framework Programme.

The governance of such a large and complex infrastructure involves over 400 project participants based in organisations countries across the length and breadth of Europe. Several bodies contribute to the overall successful running of the project through well-defined and well-organised management and control structures.

The bodies contributing to the running of the project are: the NREN *Policy Committee* (PC), with representatives from each partner; the *Executive Committee*, consisting of a small group elected by the NREN PC, the Project Management Team to discuss operational matters, and the Project Coordinator and managing partner, DANTE.

The *Policy Committee* consists of appointed representatives from each project partner. The *Executive Committee* is comprised of a small group elected by the Policy Committee. It is primarily responsible for preparing the yearly work programme for the project, and quality assurance and supervision related to its implementation. The *Project Management Team* meets each month in order to discuss operational matters. DANTE, as the coordinating partner provides overall project management and co-ordinates the project's various activities. It is responsible for financial and administrative work, including delivering progress reports to the EC.

3.2 The European and National DCI Framework

Fig. 12 – The European grid job flow (snapshot from the GridPP Real-Time Monitor, rtm.hep.ph.ic.ac.uk).

3.2.1 National Grid Initiatives (NGIs) Collaboration

The NGIs are legal national organisations responsible for:

- establishing, managing and operating the *national Grid infrastructure* to an agreed level of service and ensuring its integration with the European e-infrastructure;
- maintaining relationships with its *national stakeholders: Research Institutes* (RIs) and Research Projects carried out by *Virtual Organizations* (VO) constituted by *Research Teams* (RTs); *Resource Providers* (or centres) which offer resources to support the computing needs of the RTs in the country. The NGI and its Resource Providers form a national "business alliance" to jointly develop and "sell" a specific national marketplace solution (the national grid infrastructure) to their national researchers, each with its specific responsibilities;
- mobilizing *national funding and resources* to guarantee the sustainable availability and operation of the national grid infrastructure as required by national users and to contribute to EGI.org for carrying out common tasks;

- *representing* all its national stakeholders in the EGI Council and in relations with EGI.eu; have the capacity to sign the Statutes of EGI.org – either directly or through a legal entity representing it;
- contributing to the decisions of the EGI Council and the EGI.eu technical bodies regarding international standards and to the EGI policies and quality criteria, and ensuring adherence at the national level to the agreed criteria;
- supporting *user communities* (application independent and open to new user communities and resource providers).

Three NGIs were selected in order to better clarify the underlying collaboration model and detailed in Part II:

(C6) IGI: Italian NGI

(C7) BEgrid: Belgian NGI

(C8) BiGGrid: Dutch NGI

Key facts on NGIs brought out from the survey

NGIs are organisations set up by individual countries to manage the computing resources they provide to the European Grid Infrastructure (EGI). NGIs are EGI's main stakeholders, together with CERN and EMBL (two EIROs). Each NGI contributes a number of computing sites to the grid infrastructure. NGIs are responsible for the maintenance and running of the sites they operate. NGIs are represented on the EGI.eu Council. A survey analysis on NGIs and EIROs participating in the EGI was recently presented at the 2011 EGI User Forum, in Vilnius, Lithuania. The survey was conducted in order to directly engage the community and start a flow of communication regarding key policy issues, obtain updates on NGI legal and organisational status and understand the most pressing issues around sustainability. Legal status, number of partners, costs, funding, tax exemptions and government relations have been investigated.

The "EGI Sustainability Plan" report, (EGI.eu D2.7, 2011), presents an updated taxonomy of services and a first outline of potential business models relevant to EGI serving as the basis for future discussion and exploration. The "EGI Role Towards Europe 2020" report, (EGI.eu, 2011), provides an overview about the latest EU strategic developments relevant to the EGI community and the related recommendations. The "Alignment of EGI.eu with the ERIC Organisational Model" document describes potential advantages and disadvantages of the ERIC legal framework and stimulated a discussion within the EGI Council; the results of a related survey were presented at the EGI User Forum at Vilnius, Lithuania, reporting a significant interest for the legal instrument and the possibility to consider its adoption at the end of the EGI-InSPIRE time frame.

NGIs are organisations set up by individual countries to manage the computing resources they provide to the European e-Infrastructure (EGI). IGI, as a national grid initiative, represents a collaboration between Academic and Research Institutions and the Italian *Ministry for University and Research* (MIUR). BEgrid is a collaboration between BELNET, a public authority, and resource providers who can be categorized as national champions. BiG Grid is a collaboration between several academic and research organizations and the Netherlands Organisation for Scientific Research (through NCF).

3.2.2 European Grid Organization and related Project Collaboration

A detailed design study (EGI_DS FP7) identified processes and mechanisms for establishing a *European Grid Initiative* and defined the structure of a corresponding coordinating body. This resulted in EGI.eu being established as a foundation under Dutch law. The mandate of EGI.eu was to create and maintain a pan-European Grid Infrastructure in collaboration with the NGIs and *European Intergovernmental scientific Research Organisations* (EIROs), to guarantee the long-term availability of a generic e-infrastructure for all European research communities and their international collaborators. EGI.eu is governed by a *Council*, which has representatives from all of its participants and is responsible for providing long-term direction to the organisation, and the *Executive Board* which provides frequent guidance to the Director, who leads the organisation on a day-to-day basis.

In order to guarantee a smooth transition from the EGEE series of projects to the new EGI framework model a specific FP7-funded European project collaboration was activated, the EGI *Integrated Sustainable Pan-European Infrastructure for Researchers in Europe* (EGI-InSPIRE) project, with the goal of providing European scientists and their international partners with a sustainable, reliable e-Infrastructure that can support their needs for large-scale data analysis. EGI-InSPIRE will coordinate the transition from the EGEE project-based system to a sustainable pan-European e-Infrastructure. The four-year project will support grids of *High Performance Computing* (HPC) and *High Throughput Computing* (HTC) resources. The project is ideally placed to integrate new *Distributed Computing Infrastructures* (DCIs) such as clouds, supercomputing networks and desktop grids, to benefit user communities within the European Research Area. EGI InSPIRE follows the projects DataGrid (2002-2004), EGEE-I, -II and -III (2004-2010).

The following collaborations case were selected for analysis:

(C9) EGI.eu Organization and EGI-InSPIRE Project collaboration

Key facts on EGI brought out from the survey

The *European Grid Infrastructure* (EGI) is a federation of resource providers set up to deliver sustainable, integrated and secure computing services to European researchers and their international partners. EGI enables access to computing resources for European scientists and researchers from all fields of science, from High Energy Physics to Humanities. In EGI there are currently 13,699 users, grouped in 218 VOs. As of April 2010 15 million jobs/month were executed, distributed over 52 countries and 317 sites, using nearly 250,000 cpu cores. The scientific disciplines involved are: astronomy and astrophysics, computational chemistry, computer science and mathematics, earth sciences, fusion, HEP, Infrastructure, Life Sciences and others.

The federated resources provided by EGI are available to all scientists and researchers who are members of a *Virtual Organisation* (VO). Each VO has its own rules as to who can join their community and membership may be open to non-European collaborators. The resources coordinated by EGI are free at point of use: individual users do not have to pay to use the grid infrastructure. Participating countries and institutions contribute to the common costs of running the infrastructure. The individual resource providers (NGIs) fund and maintain the hardware in their own countries. EGI.eu is an organisation established in February 2010 to coordinate and manage the infrastructure (EGI) on behalf of its participants: NGIs and EIROs.

EGI was designed in order to add persistence and economic sustainability to the EGEE model of collaborations. After the definition of the EGI/NGIs model with the EGI Blueprint (2009) and the following establishment of EGI.eu in 2010, the National Grid Initiatives have been progressively consolidated in several countries as the structural elements behind EGI. The EGI/NGIs model borrowed several characteristics from the Géant/NRENs model, but has several distinctive elements; DCIs require *middleware* maintenance and development (and sometimes the adoption of specialized releases, as is the case of IGI) and users are organized in domain-related communities that need support in order to port their applications onto the grid or the clouds.

EGI.eu, a foundation recognised by Dutch law and headquartered in Science Park Amsterdam, the Netherlands, is an organisation established in February 2010 to coordinate and manage the infrastructure (EGI) on behalf of its participants: *National Grid Initiatives* (NGIs) and *European Intergovernmental Research Organisations* (EIROs). NGIs are EGI's main stakeholders, together with CERN and EMBL (two EIROs). Each NGI contributes a number of sites to the grid infrastructure. NGIs are responsible for the maintenance and running of the sites they operate. NGIs are represented in the EGI.eu Council. The European Project collaboration EGI-InSPIRE is a consortium of 50 partners, including 37 National Grid Initiatives (NGIs), two European International Research Organisations (EIROs) and eight partners from the Asia Pacific region.

3.2.3 Middleware Providers' Collaborations

The middleware providers' main goals are to improve the reliability, usability and stability of middleware services, intently listening to the requirements of users and infrastructure providers. Different providers (EMI, IGE, Venus-C, StratusLab, EDGI) support several middleware technologies (Arc, gLite, Unicore, Globus, Cloud Middleware, Desktop Grids Middleware, …), consolidating and improving the existing middleware services.

The following cases were selected for deeper analysis:

(C10) The EMI Grid Middleware Provider collaboration

(C11) The StratusLab OS Cloud Middleware Provider collaboration

(C12) The Venus-C Public/Private Cloud Middleware Provider collaboration

Key facts on Middleware Provider Collaborations brought out by the survey

EMI, StratusLab and Venus-C are three of the six projects funded under "Distributed Computing Infrastructures" sub-topic of the e-Infrastructures topic of the FP7 "Capacities" Specific Programme Call 7 (FP7-Infrastructures-2010-2), which closed in November 2009. The others are EGI-InSPIRE, IGE, EDGI. Apart from EGI-InSPIRE, the projects are all middleware-focused:

- the *European Middleware Initiative* (EMI) is a collaboration of the three major middleware providers in Europe, ARC, gLite and UNICORE, and other consortia. EMI aims to *deliver a consolidated set of middleware components for deployment* in EGI, PRACE and other DCIs;
- *StratusLab* will incorporate virtualisation and cloud technologies into existing and future grid infrastructures;
- *Venus-C* will develop and deploy a Cloud Computing service for research and industry communities in Europe;
- IGE, the *Initiative for Globus in Europe* provides the European link with Globus technology;
- EDGI, the *European Desktop Grid Initiative* will develop bridge middleware that integrates ARC-, gLite- and UNICORE-based Grids with Desktop Grids (BOINC and XtremWebHEP-E).

The six “DCI Projects” have come together on several occasions, in particular during the “Engaging European DCIs Together” workshop organised by the SIENA project in May-June 2010. As a result of an ongoing joint discussion, the parties have produced a written *DCI Collaborative Roadmap* that aims to harmonize these core technologies and thus strengthen and broaden the set of services underpinning all distributed computing activities.

3.2.4 Virtual Research Community (VRC) Collaboration

A *Virtual Research Community* (VRC) is defined as an organisational grouping that brings together transient *Virtual Organisations* (VOs) within a persistent and sustainable structure. A VRC must be a self-organising group that collects and represents the interests of a focussed collection of researchers across a clear and well-defined field. Named contacts are agreed upon by the VRC to perform specific roles and these then form the communication channel between the VRC and EGI (*See: MoU VRC / EGI.eu template*).

We selected in particular two relevant cases:

(C13) The WLCG HEP Physics VRC

(C14) The WeNMR VRC

Key facts on VRCs brought out by the survey

WLCG VRC is the world’s largest (grid) virtual research community. It is the outcome of a global collaboration of more than 140 computing centres in 34 countries, the four experiments at the *Large Hadron Collider* (LHC), and several national and international grid projects. The WLCG VRC mission is to build and maintain a data storage and analysis infrastructure for the entire *High Energy Physics* (HEP) community that will use the LHC at CERN. It was launched in 2002 to provide global computing resource to store, distribute and analyse the 15 Petabytes of data annually generated by LHC. The WLCG VRC runs more than 1,000,000 tasks per day on 200,000 processors from 140 computing centres in 40 countries, as well as multi GB/s data transfer on the full mesh of its tiered centres.

VRCs in other domains – especially life sciences – are steadily growing, following the example of the HEP community. WeNMR, with an initial focus in biomolecular *Nuclear Magnetic Resonance* (NMR) and *Small Angle X-ray Scattering* (SAXS) communities is the second largest VO in the Life Sciences.

3.2.5 Outreach, Dissemination, Standards and Interoperability Collaborations

The *Open Grid Forum* (OGF) & OGF-Europe are committed to driving the rapid evolution and adoption of applied distributed computing, which is critical to developing new, innovative and scalable applications and infrastructures essential to productivity in enterprise and within the science community. These goals are accomplished through open fora that build the community, explore trends, share best practices and consolidate these best practices into standards.

Within the DCI community, there is the need to accelerate and co-ordinate the adoption and evolution of interoperable DCIs. The *Standards and Interoperability for E-infrastructure implemeNtation Initiative* (SIENA) collaboration is targeting this goal through engagement with other *Standards Development Organisations* (SDOs) and major stakeholders to forge community agreements on best practices and standards for distributed computing.

SDOs relevant to RI DCIs are: *Cloud Security Alliance* (CSA), *Distributed Management Task Force* (DMTF), *Open Cloud Consortium* (OCC), *Open Cloud Computing Interface Work Group* (OCCI-WG), *Object Management Group* (OMG), *Storage Networking Industry Association* (SNIA); the Grid Technical Committee in the *European Telecommunications Standards Institute* (ETSI), the *Open Geospatial Consortium* (OGC), the *International Organization for Standardization* (ISO), the *Organization for the Advancement of Structured Information Standards* (OASIS).

As part of its International outreach programme, SIENA’s ambition is to engage major stakeholders in order to forge community agreement on best practices and standards for distributed computing. This will be operationally sought through the *Industry Expert Group* (IEG), the *Special Liaison Group* (SLG) and the *Roadmap Editorial Board* (REB) made up of international experts.

e-ScienceTalk brings the success stories of Europe’s e-infrastructure to a wider audience. The project coordinates the dissemination outputs of EGI and other European e-Infrastructure projects, ensuring their results and influence are reported in print and online. *GridCafe*, *The Digital Scientist – formerly international Science Grid This Week* (iSGTW), *GridGuide* and *e-Infrastructure Concertation Meetings* comprise its differentiated audience dissemination channels.

3.3 The (HPC) Top Parallel Computing RI Ecosystem

Fig. 13 – The PRACE European collaboration and the first tier-0 class systems: JUGENE (right), CURIE (lower left), HERMIT (upper left).

The following collaboration use cases are extracted from the PRACE project documents (D1.5, D2.1.1, D2.2.1, D2.5.1, D2.4.2).

The definition of an HPC Ecosystem can be visualised as a wide consortium including all aspects which stimulate the efficient usage of various kinds of computational resources producing high-class scientific results. The HPC service is often described by the Performance Pyramid, which consists of multiple layers:

- The European level *capability computing centres* (tier-0), which represent the highest available computing power, providing computing services to the top research groups across national borders and scientific disciplines;
- The *national and regional computing centres* (tier-1) with sufficient computing services for HPC users and to facilitate the access ramp to the resources of the European level centres;
- The *local computing centres* (tier-2) in the university environments, research labs or in other organisations;
- The *personal computer or terminal resources* (tier-3) available to individual researchers.

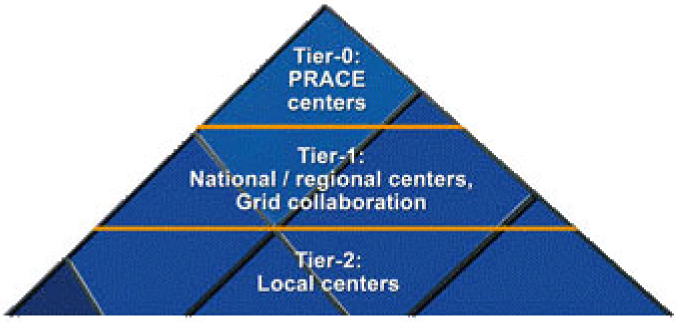

Fig. 14 – The HPC Performance Pyramid: from Tier-0 to Tier-2.

The focus of the ESFRI PRACE Project Collaboration is on the high end of computing and the enabling infrastructures for it, i.e. the tier-0 and the HPC parts of tier-1. PRACE will also coordinate and collaborate with other projects in those domains, most notably the DEISA 2 (tier-1) and EGI/EGEE (tier-1, tier-2 and tier-3). In the following a brief description of the HPC ecosystem building blocks will be given.

3.3.1 Tier-0 Capability Computing Centres Collaborations

The Principal Partners Committee (PPC, called hosting members in the PRACE AISBL) decides on tasks that are essential for the definition, hosting, construction, funding or financing of the Tier-0 HPC infrastructure. The MB takes all other decisions of strategic importance. PPC and MB have face-to-face meetings every 2 months. These project meetings are organised jointly with PPC and MB meetings of the PRACE Initiative.

3.3.2 Tier-1 National and Regional Computing Collaborations

Tier-1 *General Partners* are essential to achieving the goals of developing HPC in Europe. Tier-0 and Tier-1 are both essential components of the constituency, and are extremely important contributors both with respect to HPC operation and procurement and in the area of leading HPC application in science and technology. The contributions of the General Partners may vary, depending on the funding model chosen, with some general partners paying more into the budget than others and some consideration of the financial contribution given being reflected in voting rights.

We selected the following National Computing Centre collaboration cases for a deeper analysis:

(C15) The Netherlands Computing Facilities Foundation (NCF)

(C16) The Finland CSC HPC Centre

Key facts on HPC Centres brought out by the survey

The PRACE "Analysis of HPC Ecosystem" (D2.5.1) reports the results of an analysis of an assessment of collaboration and integration opportunities for PRACE. The HPC Centres analysed have budgets in the order of 15-25 millions of euro, mainly from Public Authorities.

NCF is part of the *Netherlands Organisation for Scientific Research* (NWO). NWO stimulates and finances research in every conceivable scientific discipline and facilitates innovations. NWO is autonomous but falls under the responsibility of the Ministry of Education, Culture and Research. CSC – IT Centre for Science Ltd (CSC) is the national organization responsible for academic supercomputing, research networking, data storage and scientific software and database services. It is a non-profit company whose shareholder is the Ministry of Education and Culture.

3.3.3 The PRACE Collaboration

Building a world-class pan-European High Performance Computing (HPC) Service is a highly ambitious undertaking that involves governments, funding agencies, centres capable of hosting and managing the supercomputers, and the scientific and industrial user communities with leading edge computationally needy applications.

In contrast to Research Infrastructures that focus on a single scientific instrument, an HPC Infrastructure has two unique characteristics: supercomputers serve all scientific disciplines and tier-0 supercomputers have a three year depreciation cycle as tier-0 implies continuous leading edge services. This requires a periodic renewal of the systems and a continuous upgrade of the infrastructure. Furthermore, novel architectures and system designs will be created by vendors for leadership systems. There will be several different systems each of them best serving a particular application spectrum.

This fact mandates a distributed Research Infrastructure (RI), since no single site can host all the necessary systems because of floor space, power, and cooling demands.

The *PRACE organisation* is responsible for providing access to a suite of Tier-0 systems for use by European research communities. PRACE will be able to:

- define and implement a strategy for providing a world class HPC infrastructure in Europe,
- manage the formation of a suite of complementary Tier-0 systems in Europe,
- manage the contributions of partners,
- operate an open and fair access system based on peer review to the Tier-0 services,
- manage the interaction with industrial organisations wishing to access PRACE systems,
- perform training and computational science R&D activities,
- interact with multiple stakeholders in order to promote HPC in Europe with a long term and sustainable approach,
- deliver appropriate accounting, administration, human resources, marketing and communication activities,
- provide a secretariat to PRACE governance bodies.

We selected the PRACE collaboration model as an important case to study:

(C17) The PRACE HPC Centres Collaboration

Key facts on PRACE brought out by the survey

As reported in the PRACE D2.5.1 "Analysis of the HPC Ecosystem", successful collaboration to utilize petascale HPC services and resources requires major economic and human investments from a variety of different sources and stakeholders. To this extent a collaborative effort towards a highly sustained capability computing infrastructure was brought forward with PRACE. The project focused on scalable code development, integration with national infrastructures, sufficient data repositories, high capacity networking based on global standardisation efforts and education and training efforts.

The PRACE RI was created as an international non-profit association on April 23, 2010 by 19 members, representing Austria, Bulgaria, Cyprus, Czech Republic, Finland, France, Germany, Greece, Ireland, Italy, The Netherlands, Norway, Poland, Portugal, Serbia, Spain, Sweden, Switzerland, Turkey and the UK. Additional European states are invited to join. The seat of the PRACE AISBL is in Brussels.

The PRACE Research Infrastructure consists of several *supercomputers* in the *Petaflop/s range*. The current systems are:

- **JUGENE**, *1 Petaflop/s*, an IBM BlueGene/P at Forschungszentrum Jülich, Germany,
- **CURIE**, *1.6 Petaflop/s* (when the last phase of the installation will be completed in October 2011), a Bull at CEA, France.

A third system will be installed in the autumn of 2011:

- **HERMIT**, *1 Petaflop/s,* a Cray XE6 at the Computing Centre of the University of Stuttgart.

Also, will become operational in 2012:

- **SuperMUC**, *3 Petaflop/s*, an IBM at Leibniz Rechenzentrum, Münich, Germany.

The *PRACE organisation* is responsible for providing access to a suite of Tier-0 systems for use by European research communities. PRACE is a cooperation of 21 partner institutions of 21 European countries. Most of them represent the national HPC service providers.

3.3.4 Hardware Vendors and Industry Collaborations for Innovation

PRACE and hardware vendors collaborate in various areas including prototyping activity, application benchmarking and other technical development. The permanent research platform *Advanced HPC Technology Platform* (AHTP) is being implemented in PRACE. This includes the designation of the partners of AHTP from PRACE partners and collaborating industrial partners. AHTP will actively seek cooperation with European projects with the potential of contributing to petascale computer and communication components and supporting similar activities in order to derive benefit from all experiences and innovative ideas available throughout Europe. PRACE expects the vendors to be able to share technical and other relevant information, which might be helpful in developing the PRACE collaboration or services in the current computing environment. The actions that shareholders need to take include active participation in various European HPC forums, proactive interest towards PRACE work and open attitude towards new ideas on how to develop HPC activities in Europe. PRACE will also collaborate with recently established HPC collaborations targeting European industry, such as PROSPECT and TAL.

The *PRACE advisory group for STRAategic TechnOlogieS* (STRATOS) has been created – as a complement to the work that is aimed at the procurement, construction and operation phases of HPC services – to foster the development of components and technologies for future multi-petascale systems. Within STRATOS, partners from PRACE and industrial consortia, including more than eighty organisations, co-operate on the specification and developments of such components. The STRATOS MoU was signed in December 2008 by twelve PRACE partners and the consortia PROSPECT and Ter@Tec. Its long-term work plan comprises joint activities on Exascale Software, and Green-IT and HPC Leadership resources.

3.4 The MNT Collaboration Facilities Interchange RI Framework

Fig. 15 – European MNT facilities: IMEC (upper left), LETI (upper right), Fraunhofer IISB (lower left), TNI (lower right).

3.4.1 MNT National Research Centres / Facilities

One of the earliest outputs of many MNT national and regional technology policies was the development of MNT centres (MNTC). The authors [Kautta, Walshb, Bittnera, 2007] reviewed the variety of forms and defined the nature of many of these major MNTCs from around the world by describing some of their similarities and differences. MNT National Centres – called *Research Organizations* (ROs) in the PRINS documents – are the building blocks of wider forms of collaborations that combine complementary MNT centres in a more capable MNT Research Infrastructure.

We identified a set of MNT National Centres to analyze:

(C18) The Irish Tyndall National Research Centre

(C19) The Belgian IMEC National Research Centre

Key facts on MNT National Centres brought out by the survey

The Irish NAP is based on the European model of providing fully funded trans-national access to key research infrastructure. The current model of fully funded access often works very well as it offers the users a straight-forward route to access. The disadvantage of the model is that it relies on continuous funding from the funding agencies which raises a question about long-term sustainability. The *National Access Programme* (NAP) offers researchers in Ireland access to the facilities and expertise available at the Tyndall National Institute (Tyndall). The funding for NAP comes from a PA (SFI) and is distributed by an NC (Tyndall). SFI regularly reviews the activities of Tyndall in relation to how the funds are allocated.

IMEC vzw is a non-profit organization under Belgian law. It has sister companies in The Netherlands, Taiwan and China, and representation offices in the US and in Japan. IMEC vzw receives an annual grant from the Government of Flanders. In contrast with the NAP case, only about 16% of the budget of IMEC is an annual grant from its Public Authority (the Government of Flanders). The other part of the budget is based research financed by public co-funded programs (National research funds, European Commission, European Space Agency) and R&D performed with industrial partners from all over the world. Its staff of more than 1,750 people include over 550 industrial residents and guest researchers. The multi-cultural research environment constitutes researchers and students from more than 60 different nationalities.

3.4.2 MNT ROs Alliances and Collaborations for Transnational Access

A combined Research Infrastructure of the MNT ROs has the capability to offer the following access mechanisms for academia (*see: PRINS Final Report*):

- *hosting research teams* in the RI and/or building up common R&D laboratories that can enhance the long term collaboration between Academia and the hosting Research Organization;
- favouring better *mobility for researchers*, essential to leverage the impact of the research community and enhance the industry / research centre / academia linkages;
- allowing *effective collaboration between Academia and the MNT RI* by cooperating in scientific and technical projects;
- providing academic research teams with *basic materials/data for their own research programs in ways that are complementary to the direction being followed by the RI.* RI will *leverage the whole investment* as well as supporting the knowledge base and establishing *new technology directions* to pursue for the RI.

The access given by the combined RI could allow the academic community to perform advanced research in several domains and to remain at the leading edge of the research world and increase the industrial relevance of the research, thereby lowering the gap between academic research and industrial R&D.

A collaboration model can also be developed through a strategic Alliance between the major European MNT Research Centres by creating a common technology portfolio in the form of a distributed platform for research and development. For this case, the following networking activities were identified (*see: MNTEurope DoW*):

- creation of the *environment and legal framework* to allow the Integration of Activities to lower the barriers to collaboration between the participant organisations, in particular legal issues and those related to the lack of personal relationships will be targeted for improvement,
- *preparation of the integration* to build a common view on future developments;
- organization of the *operational phase* of the alliance: global management, external collaboration, dissemination of the alliance concept.

The integration of MNT Centres in a Research Infrastructure can also benefit from a competitive access model similar to the HPC one, giving new opportunities to the best research teams via a peer reviewed access model. Costs for access are then covered by the EC. The Transnational peer-reviewed model has the following characteristics (see the ANNA I3[10]):

- *EC-funded* transnational access to infrastructure (laboratories, clean room, meteorology) at locations,
- access to instrumentation and analytical services is either *in person* ("hands-on") or *remotely* by suitable (electronic) communications,
- potential users of the infrastructure are researchers or groups of researchers from small & medium enterprises, large scale industry, research centres, or universities,
- interested users *apply for research access* by submitting a short project proposal,
- selection of user proposals is by "peer review" on the basis of scientific/technical merit.

Each installation/facility is listed, described, and a temporal unit of access is specified (e.g., week, day). The installations are managed by the cooperating organizations that own them.

The following cases of multi-MNTC collaborations were selected:

(C20) The Epixnet Network of Excellence

(C21) The MNT Europe Project Collaboration

(C22) The Sinano Institute Collaboration

(C23) The MNT Heterogeneous Technology Alliance (HTA)

Key facts on multi-MNT collaborations brought out by the survey

ePIXfab is a collaboration between IMEC and CEA-LETI, with a collaboration agreement between the two institutes. ePIXfab is funded by the European Union, through PhotonFAB. The funding covers the operational expenses for European users but not materials costs. All

[10] See*:* www.i3-anna.org/context.jsp?ID_LINK=2&area=8.

process costs are fully paid by the users. The cost in a run is shared between the users, which makes the expensive technology affordable for universities and research institutes.

MNT Europe and its extension were aimed at increasing the cooperation between five major European Research Institutes in the field of micro- and nanotechnologies: CEA-Leti, CSEM, FhG IISB and IZM, IMEC, and Tyndall Institute. It was an EC-funded FP6 I3 project.

Four microsystems partners (CEA-Léti-Liten, Fraunhofer Gesellschaft Microelectronic Alliance, VTT, and CSEM) have signed an agreement to cooperate with their MNT research infrastructures in the MNT Heterogeneous Technology Alliance in a purely private effort.

The SiNANO Institute is the largest MNT collaboration surveyed, as it gathers twenty Academic Institutions from eleven European countries: Grenoble INP (FR), Chalmers University (SE), IUNET (IT), KTH Royal Institute of Technology (SE), NCSR Demokritos (GR), Newcastle University (UK), Research Centre Jülich (DE), Tyndall Institute (IE), Université Catholique de Louvain (BE), Université de Lille (FR), University of Glasgow (UK), University of Liverpool (UK), University of Warwick (UK), University Rovira i Virgili / University of Granada (ES), Uppsala University (SE), Warsaw University of Technology (PL), Montpellier University (FR), University of Twente (NL), ICN (ES), ITE (PL). The funding model is: 30%-70% from National Funding, 70%-30% from European + National + Industrial Projects funding.

3.5 The Research Data Infrastructure Framework

Fig. 16 – ESO Science Archive Data (image from www.eso.org).

As noted in report “Riding the wave” [High Level Expert Group on Scientific Data, 2010], the “virtual lab” is already real, with the ability to undertake experiments on large instruments in other continents remotely in real time. Researchers with widely different backgrounds – from the humanities and social sciences to the physical, biological and engineering sciences – can collaborate on the same set of data from different perspectives.

There is a need for a *Collaborative Data Infrastructure* that would allow different companies, institutes, universities, governments and individuals to share data and inter-operate seamlessly. The emerging infrastructure for scientific data should be flexible but reliable, secure yet open, local and global, affordable yet high-performance.

RDIs can be classified according to the middleware level listed in the "e-IRG Report on Data Management" [e-IRG-DMTF, 2009]:

- *Data Storage*, for interoperable storage;
- *File Catalogues and persistent identifiers*, to manage files within an information environment and to maintain interoperability over long periods of time;
- *Metadata Catalogues*, for indexing and describing the resources and their quality;
- *Federated Databases*, to transparently integrate heterogeneous database systems;
- *Digital Repositories*, to manage digital objects, where an object is anything from a simple file to a complex object composed of various files, metadata, and relations with other objects and functionalities.

The OSIRIS project looked mainly at this last category, considering separately general infrastructures for research (following) from domain-specific infrastructures (subsequent paragraph).

3.5.1 Digital Libraries and Infrastructures for Research in Europe

As reported[11] in the Communication from the Commission of 30 September 2005 to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – i2010: digital libraries [COM(2005) 465 final – Official Journal C 49 of 28.2.2008], "Digital libraries are *organised collections of digital content made available to the public*." The content is material that *has been digitised*, such as digital copies of books and other 'physical' material from libraries and archives. Alternatively, they can be based on information *originally produced in digital format*. This is increasingly the case in the area of scientific information, where digital publications and enormous quantities of raw data and software are stored in digital repositories.

The i2010 *Digital Libraries initiative* covers both *digitised* and *born digital* material. Three areas have been identified: (a) *Digitisation* of Analogue Collections, (b) *Online Accessibility* of Digital Content, (c) Digital *Preservation*. The identified key issues for the *European Digital Library* are:

- **Choice of materials:** the selection process should be 'bottom up'. Libraries and archives which own the materials should decide what is included as part of their own strategic development plans and available resources.

[11] See: http://europa.eu/legislation_summaries/information_society/l24226i_en.htm.

- **Meeting user needs:** the needs of the user should be central. Development will be demand-driven, but it is also important to take a longer term and visionary view of what the user will get from the library in the way of services.
- **Leveraging public/private partnerships:** at present, digitisation efforts in the Member States are progressing rather slowly. Public/private partnerships or private sponsorship can be a useful means to complement public funding and help accelerate digitisation.
- **European value:** financing digitisation is mainly a responsibility of the Member States. EU programmes can contribute where there is most European added value, for example by helping to aggregate digitised material across borders.
- **Intellectual property:** under current legislation only public domain works (where there is no longer copyright) can be made freely available to the public online. Hence, only works with specific rights clearances will be accessible through a digital library.

Europeana is a simple but powerful tool for finding resources from all over Europe. Books, journals, films, maps, photos, music etc. will be available for everyone to consult – and to use, copyright permitting. For example, the library will be a rich source of materials for the creative and information industries in developing new products and services, for tourism and for teaching.

Some 2.5 million research articles are published in 25,000 peer-reviewed journals and conference proceedings worldwide every year. Currently, just 15%-20% of these articles are available in Open Access repositories or Open Access journals. The rest are only accessible through pay per read schemes or by paying for a subscription to the publication. The *Open Access Infrastructure for Research in Europe* (OpenAIRE), launched December, 2nd 2010 by the EC at the University of Ghent in Belgium, could eventually *open up access* to all scientific papers and data produced by researchers funded by the EU's FP7, including scientists receiving grants through the *European Research Council* (ERC), and beyond. OpenAIRE will provide *a network of open repositories providing free online access to knowledge produced by scientists receiving grants from the FP7 and ERC*, especially in the fields of health, energy, environment, parts of Information & Communication Technology and research infrastructures, social sciences, humanities and science in society. Since FP7 started in 2007, some 10,000 projects have been funded. The OpenAIRE project could also lead to new ways of *indexing*, *annotating*, *ordering* and *linking* research results – and new methods to automate all this. This could trigger the development of new services in addition to the information infrastructure which OpenAIRE provides. This project runs a helpdesk in each of the 27 European countries, consisting of a network of experts and a portal of tools helping researchers to make their articles available online.

We selected for our inventory of cases:

(C24) The Open Access Infrastructure for Research in Europe

(C25) The EUDAT (EUropean DATa) Project

Key facts on RDI infrastructural collaborations brought out by the survey

OpenAIRE aims to support the implementation of Open Access in Europe. It provides the means to promote and realize the widespread adoption of the *Open Access Policy*, as set out by the ERC Scientific Council Guidelines for Open Access and the Open Access pilot launched by the European Commission. OpenAIRE objectives are to: (1) build *support structures* for researchers in depositing FP7 research publications, (2) establish and operate an *electronic infrastructure* for handling peer-reviewed articles, (3) explore the *requirements, practices, technologies*, to deposit, access and manipulate *research datasets*. OpenAIRE is a FP7 CPCSA three-year project. The *National Open Access Desks* connect researchers, research institutions, and policy makers at a national level on the one hand, and the OpenAIRE project services on the other. The OpenAIRE network of Open Access desks is structured similarly to the Europe-wide information network on European Research Programmes.

EUDAT is a project to create a *collaborative scientific data infrastructure* with the capacity and capability for meeting future researchers' needs in a flexible and sustainable way, across geographical and disciplinary boundaries. EUDAT will undertake the first comprehensive European review of the approaches to deployment and use of a common and *persistent* data e-Infrastructure, the services being built and delivered on top of this infrastructure and the limitations of its services. EUDAT will be ensuring the *authenticity*, *integrity*, *retention* and *preservation* of data deposited by its users, especially those marked for long term archiving.

The EUDAT consortium brings together e-infrastructure providers, research infrastructure operators, specialist application/solution providers and researchers from a range of scientific disciplines under several of the main ESFRI themes. The depth and breadth of coverage allows a collaboration between for example, data centres such as JUELICH, KIT, CSC, BSC, SIGMA, SARA and SNIC, that support multiple customer groups as well as three infrastructure-providers (CSC, PSNC and SIGMA) operating 'end-to-end' across the networking, grid, HPC and data-management layers of the e-Infrastructure. Equally important, given the goal of self-sustaining operation within three years, is the involvement of two agencies with national funding responsibilities (SNIC, STFC) able to litmus test the proposition as it emerges.

3.5.2 Scientific Repositories collaborations

Projects relevant to scientific e-Infrastructures are:

- DRIVER II (e-Infrastructures and info-structures) – delivers a pan-European infrastructure *federating scientific repositories*. It uses open standards and supports complex information objects. Digital repositories will make both published and experimental data widely available for use and thus support new paradigms for research.
- PARSE.Insight aims to *help define the infrastructure needed to preserve and use the digitally encoded information* on which our society increasingly depends and which future generations will inherit.

Specific domain projects are also relevant e-Science use cases:

- EURO-VO-AIDA (astronomy): moving the astronomical European Virtual Observatory into a fully functioning operational phase,
- GENESI-DR (Earth observation): open and seamless access to Earth science repositories (space, airborne and in-situ sensors data),
- IMPACT (bio-informatics): improving protein annotation through coordination and integration of databases,
- METAFOR (climatology): common information model and tools for using climate data and models,
- NMDB (space physics): digital repository for cosmic ray data, pooling archives and collecting observations real-time,
- PESI (biodiversity): taxonomically validated standardised nomenclatures for biological and biodiversity management.

In the following, we recall the results of the survey of data initiatives conducted in [DMTF, 2008], 1.3, pp 16-50.

3.5.3 Arts and humanities, social sciences

In social sciences, important datasets are often collected not by research teams but by government departments to inform policy makers. These data are usually of high quality, are national samples or census data, and are often under-utilised with high secondary value as a source of information for social science research. Social science archives recognised this potential early and some have been able to negotiate access to these data for the wider research community.

Data Initiatives: CASPAR – *Cultural, Artistic and Scientific Knowledge for Preservation, Access and Retrieval*, CLARIN – *Common Language Resources and Technology Infrastructure*, TextGrid, IASSIST – *International Association for Social Science Information Service and Technology*, DOBES – *Dokumentation Bedrohter Sprachen, HRELP-Hans Raising Endangered Languages Project*, **DARIAH** – *Digital Research Infrastructure for the Arts and Humanities*, **DC-NET** – *Digital Cultural Heritage NETwork,* CESSDA – *Council of European Social Science Data Archives, DANS -Data Archiving and Networked Services*, UKDA – *UK Data Archive, University of Essex*, ESDS – *Economic and Social Data Service*, RELU-DSS – *Rural Economy and Land Use Data Support Service*, SHARE – *Survey of Health, Ageing and Retirement in Europe*, COMPARE, DCC – *Digital Curation Centre*, DPC – *Digital Preservation Coalition*, ICPSR – *Inter-University Consortium for Political and Social Research*, NCeSS – *National Centre for e-Social Science*.

Of these, two were selected for further elaboration.

(C26) The Digital Cultural Heritage Network (DC.NET)

(C27) DARIAH ESFRI

3.5.4 Health Sciences

The European Union's Member States are committed to sharing their best practices and experiences to create a European e-Health Area, thereby improving access to and quality of health care at the same time as stimulating growth in this industrial sector. The *European e-Health Action Plan* plays a fundamental role in the European Union's strategy. Work on this initiative involves a collaborative approach among several departments of the Commission services. The European Institute for Health Records is involved in the promotion of high quality electronic health record systems in the European Union.

Data Initiatives: ELIXIR – *European Life-Science Infrastructure for Biological Information*, EATRIS – *European Advanced Translational Research Infrastructure in Medicine*, ECRIN – *European Clinical Research Infrastructures Network*, EU-OPENSCREEN – *European Infrastructure of Open Screening Platforms for Chemical Biology*, INFRAFRONTIER – *European Infrastructure for Phenotyping and Archiving of Model Mammalian Genomes*, BBMRI – *Biobanking and Biomolecular Resources Research Infrastructure*, EBI – *European Bioinformatics Institute*, BIOSAPIENS, EMBRACE – *European Model for Bioinformatics Research and Community Education*, EMMA – *European Mouse Mutant Archive*, EUMODIC – *European Mouse Disease Clinic*, HEALTH-E-CHILD.

3.5.5 Natural Sciences and Engineering

There are quite different domain-specific practices in the natural sciences. Astronomy and earth and environmental sciences are well-organised communities with well-established best practices for the preservation of data. This is not the case for the High Energy Physics community despite on-going efforts around the upcoming LHC. A nascent initiative around neutron and photon sources in Europe has initiated discussions on *how to preserve the data* originating from the increasing number of these large facilities for multi-disciplinary research. This initiative will seek to address the need to *preserve and structure access to the data* generated by a community of more than 25,000 scientists in Europe.

Data Initiatives: BODC – *British Oceanographic Data Centre*, OpenDOAR – *The Directory of Open Access Repositories*, DRIVER-II – *Digital Repository Infrastructure Vision for European Research*, METAFOR – *Common Metadata for Climate Modelling Digital Repositories*, D4SCIENCE – *Distributed Collaboratories Infrastructure on Grid Enabled Technology for Science*, GMES – *Global Monitoring for Environment and Security*, HMA – *Heterogeneous Missions Accessibility*, GEOLAND – *Integrated GMES Project on Land Cover and Vegetation*, MyOcean – *Ocean Monitoring and Forecasting*, INSEA – *Data Integration System for Eutrophication Assessment in Coastal Waters*, MERSEA – *Marine EnviRonment and Security for the European Area*, PARSE.Insight – *Permanent Access to the Records of Science in Europe*, SOSI – *Spatial Observation Services and Infrastructure*, CLIMATE-G, DEGREE – *Dissemination and Exploitation of GRids in Earth sciencE*, SeaDataNet, EMODNET – *European Marine Observation and Data Network*, HIDDRA – *Highly Independent Data Distribution and Retrieval Architecture*, GENESI-DR – *Ground European Network for Earth Science Interoperations*, **LIFEWATCH**, BIOCASE, GBIF – *Global Biodiversity Information Facility*, EDIT – *European Distributed Institute of Taxonomy*, ENBI – *European Network for Biodiversity Information*, MARBEF – *Marine Biodiversity and Ecosystem Functioning*, MARINE GENOMICS EUROPE, SYNTHESYS – *Synthesis of Systematic Resources*, APA – *Alliance for Permanent Access*, DELOS, PDB –

Protein Data Bank, EuroVO-AIDA – *European Virtual Observatory-Astronomical Infrastructure for Data Access*, HEP – *High Energy Physics*, ICOS – *Integrated Carbon Observation System*.

We selected:

(C28) Lifewatch – Italian National Network

Key facts on RDI domain-related collaborations brought out by the survey

Several efforts are currently working towards the goal of digital infrastructures servicing specific fields of research, like Humanities (for DC.NET and the DARIAH ESFRI) or Natural Sciences (as is the case with Lifewatch).

DC-NET, an ERA-NET project in the 7th framework programme of the EC, is checking how existing e-Infrastructures can be used in the domain of digital cultural heritage and will define priorities in their research topics to maximise the possible advantage of the new perspectives that are created by using national and international e-Infrastructures. The mission of DARIAH is to enhance and support digitally-enabled research across the humanities and arts. The structure of DARIAH is based on the collaboration of different scientific communities in the Arts and Humanities of several European countries.

Lifewatch is a distributed RI that has as a goal to integrate data on environment and biodiversity from other RI (GBIF, National CHM, GEOSS, ..) and to offer a place to access an integrated virtual environment to share applications, workflow for integrated analysis powered by in house computing farms and computing system put forward by other RIs (i.e. EGI, PRACE). LifeWatch will be an intergovernmental organization under the jurisdiction of the European Commission. Membesr of the organization include the state Public Administrations (PA) that are interested in the goal of LifeWatch. But, to work correctly LifeWatch will need to involve the scientific community. Within LW, the scientific community is represented by a National Champion that is a member of a National Network. Depending on the different countries, the option to participate in LW was promoted by PA or NC. The actual type of collaboration varied greatly among member states.

3.6 The Remote Instruments access model

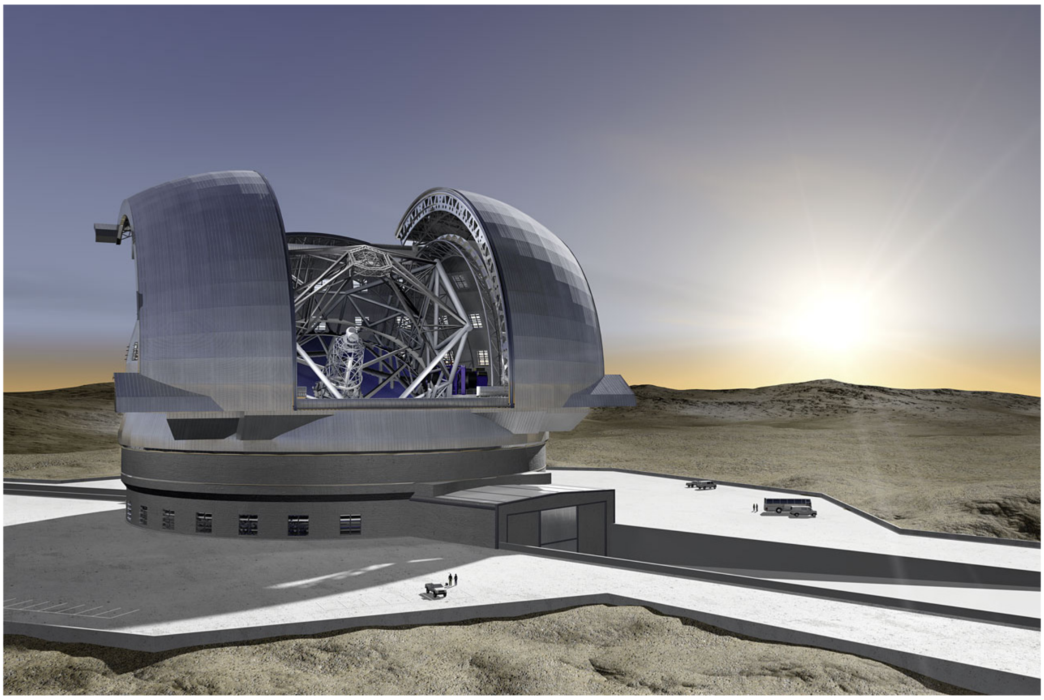

Fig. 17 - Artist's impression of the European Extremely Large Telescope (image from www.eso.org).

Coordinated and secure access to instruments, data and computational resources is an important requirement for the effective remote usage of these instruments by applications and their users. To fulfil these requirements, several experiences (like the DORII project and, in a different context, the CYCLOPS project) made use of existing e-Infrastructures in Europe – the EG infrastructure – adding the necessary components and services to facilitate remote instrumentation.

As noted in the ESFRI Implementation Report [ESFRI, 2009], Research Infrastructures both *produce* and *distribute* huge volumes of data. The attention of ESFRI has been more focused on *Digital Repositories* [DMTF, 2009] than on an *Instrument Access model*, which is still in a defining phase. There is the need for the development and continuous improvement of the underlying (data) e-infrastructure, allowing *data acquisition, transfer and analysis*, as well as *data conservation and administration*. As long as the ESFRI roadmap advances, however, the necessity of a remote access model to instruments will grow and the related scientists' requirements be clarified, bringing a more precise definition of the model and the needed e-infrastructure behind it.

3.6.1 The Instrument Element infrastructure model

(*See* www.dorii.eu) The *Deployment of Remote Instrumentation Infrastructure* (DORII) project aimed at deploying an e-Infrastructure for remote instrumentation. DORII applications require the integration of scientific instruments with computational and storage resources to facilitate data acquisition, storage and processing. Coordinated and secure access to instruments, data and computational resources is an important requirement for the effective remote usage of these instruments by applications and their users. To fulfil the above requirements the DORII project *utilized the capabilities of Grid infrastructures*.

The main principle is *the use of existing e-Infrastructures in Europe adding the necessary components and services to facilitate remote instrumentation*. The DORII eInfrastructure is mainly based on the EGI/EGEE infrastructure and its middleware of choice gLite (see: EMI gLite). The middleware service dealing with the management of remote instrumentation is the *Instrument Element* (IE) that was built by the DORII project. To deal with the interactivity requirements of the applications, the DORII eInfrastructure deployed a selection of services built by the Interactive European Grid Project (int.eu.grid).

The first version of the DORII e-Infrastructure comprised *resource centres* (sites) distributed among the partners of the project in several countries such as Germany, Greece, Italy, Poland and Spain. Ten resource centres were already available to the DORII infrastructure in its first version, providing more than 2300 non-dedicated CPUs and several terabytes of storage.

(C29) The Instrument Element (IE) Infrastructure Access Model

3.6.2 The ESFRI instrument access model

As noted in the ESFRI Implementation Report [ESFRI, 2009], Research Infrastructures produce and distribute huge volumes of data. The open access to the data generated, as well as the best use of the infrastructures themselves, require the development and continuous improvement of the underlying e-infrastructure, allowing *data acquisition, transfer and analysis*, as well as *data conservation and administration* to make both data and infrastructure easily accessible to scientists. The attention of ESFRI has however been more focused on *Digital Repositories* [DMTF, 2009] than on an *Instrument Access model*, which is still in a defining phase.

3.6.3 The Global Monitoring access model

(*See: CYCLOPS project*) CYCLOPS brings together two important Communities: *Global Monitoring for Environment and Security* (GMES) and GRID, focusing on the operative sector and needs of European Civil Protection (CP). The main objectives of CYCLOPS are:

- To disseminate EGI results to the CP Community, assessing EGI infrastructure for CP applications. A variety of activities will focus on dissemination and outreach, training, workshops, possibly in close relation with EGI events and on promoting a close collaboration between the two communities.
- To provide the EGI Community with knowledge and requirements that characterise CP services. These requirements will also be used to assess the possibility for the development of an advanced grid platform enabling Real Time and near-Real Time services and the implementation of a security infrastructure very close to the defence systems standards.
- To evaluate the possibility of utilising the present EGI services for CP applications and developing the research strategies to enhance the EGI platform.

- To develop the research strategies to enhance the EGI platform, especially for Earth sciences resources. CYCLOPS contributed to the EU policy developments establishing liaisons and synergies with other existing projects and initiatives dealing with GMES, GRID and complementary sectors, among them: PREVIEW, Risk EOS, RISK-AWARE, BOSS4GMES, EGI Networking Activities and Application Support, e-IRG and INSPIRE.

(C30) The Global Monitoring Access Model (Cyclops)

Key facts on Remote Access model brought out by the survey

Grids are composed mainly of computing and storage resources. The *Instrument Element* (IE) component adds a possibility to include such devices as scientific instruments and sensors in the data elaboration process, providing users with a simple way of attaching their scientific instrumentation to gLite-based grids. The *Access Model* behind the IE would allow the effective *connection of scientific instruments* to the computation capabilities of the DCI.

From a more organisational point of view, the *CYber-Infrastructure for CiviL protection Operative ProcedureS* (Cyclops) is one of the first projects in Europe to include the R&D and *Civil Protection* (CP) communities in order to develop a proof of concept in using a computational grid infrastructure to build real time or near real time CP applications. The *Global Monitoring for Environment and Security* (GMES) concept was endorsed by the EU Commission in 2001 with the aim of "establishing, by 2008, a European capacity for Global Monitoring of Environment and Security" to gather and use all available data and information in support of sustainable development policies.

3.7 The Future Internet service-oriented vision

Fig. 18 – Image from www.enisa.europa.eu.

As reported in [WGFI, 2008], a multiplicity of new usage patterns and a plethora of requirements not foreseen when the Internet was designed, justifies a fresh look at the Internet architecture. Some of the key emerging usage trends are outlined:

– *Mobility* [EU projects: eMobility, FIRE],
– *End to End very high rate throughput* (broadband) [EU projects: eMobility, NEM, ISI, FIRE, Photonics21],
– *Security and Trust, Privacy* [EU projects: eMobility, NEM, EPoSS, NESSI, RISEPTIS Think Tank, ISI],
– *Device connectivity*, *coupling of virtual world data with physical world information* (RFID, sensors) [EU projects: eMobility, EPoSS, NESSI, FIRE],
– *User generated services*, as a follow up to *user generated content* [EU projects: NEM, NESSI],
– *3D becoming mainstream* [EU projects: NEM],
– *Negotiated management and control of resources, negotiated SLA's* [EU projects: eMobility, NEM, NESSI],
– *User controlled infrastructure*.

Several EU Member States have also initiated research activities that are directly or indirectly related to the Future of the Internet. These national activities are complemented by EU level activities, and notably through a set of about 80 FI related projects that have emerged as a result of the first calls of the ICT Thematic priority of FP7.

3.7.1 The FP FIRE Initiative[12]

The *Future Internet Research and Experimentation* (FIRE) Initiative is creating a *multidisciplinary research environment for investigating and experimentally validating highly innovative and revolutionary ideas for new networking and service paradigms*. FIRE is promoting the concept of experimentally-driven research, combining visionary academic research with the wide-scale testing and experimentation that is required for industry. FIRE works to create a *dynamic, sustainable, large scale European Experimental Facility*, which is constructed by *gradually connecting and federating existing and upcoming test-beds* for Future Internet technologies.

Ultimately, FIRE aims to provide *a framework in which European research on Future Internet can flourish* and establish Europe as a key player in defining Future Internet concepts globally. With a strong network focus, the first wave of FIRE projects was launched in summer 2008, with a budget of 40 million Euros and each year sees this increasing along with the scope of the projects.

We selected FIRE as the relevant Future Internet case:

(C31) The FP7 FIRE Initiative

[12] See: http://cordis.europa.eu/fp7/ict/fire.

Key facts on Future Internet brought by the survey

To carry out experimentally-driven research on the Future Internet, researchers need a testing facility to validate their analysis of networks, service architectures and paradigms. An experimental facility on Future Internet technologies must broadly support research on networks and services, in order to compare current and future approaches. Practical experiments are needed to give credibility and raise the level of confidence in the research finding. Furthermore, the experimentation must be performed on a large scale to be representative, convincing, and to prove the scalability of the tested solution.

FIRE provides a *research environment* – called FIRE Facility – for investigating and validating highly innovative and revolutionary ideas. It consists of *several FIRE projects* that are developing an experimental facility for open use. Each FIRE project develops a large-scale testbed or a federation of testbeds, which they contribute to the common FIRE Facility offering. The first-wave FIRE Facility projects are OneLab2, PII, FEDERICA, WISEBED and Vital++. These are the pillars of the FIRE facility, forming the framework on which the second wave of FIRE facility projects will be built. The second wave includes projects such as BonFIRE, CREW, OFELIA, SmartSantander and TEFIS.

The FIRE initiative is part of the FP7 program. Integrated Projects (IP) and will run for two to three years in order to develop experimental infrastructure and to open it up to potential users. There are no direct links to local *Public Authorities*. The *National Champions* may be part of the projects and may bring in their infrastructure to be further developed and used. An example is the iLab.t[13] of IBBT that is used in three IP projects (Crew, Ofelia and Bonfire). Another example is Grid500 in France[14].

3.8 Joint Efforts and Projects

For completeness, the following joint efforts towards better coordination of existing e-Infrastructures are also described.

3.8.1 Joint Efforts

European Reflection Group (e-IRG); web site: www.e-irg.eu

The *e-Infrastructure Reflection Group* was founded to define and recommend best practices for pan-European electronic infrastructure efforts. It consists of official government delegates from all the EU countries. The e-IRG produces white papers, roadmaps and recommendations, and analyses the future foundations of the European Knowledge Society.

European e-Infrastructure Forum (EEF): web site: www.einfrastructure-forum.eu

The *European e-Infrastructure Forum* is a forum for the discussion of principles and practices to create synergies for distributed Infrastructures. The goal of the European e-Infrastructure Forum is the achievement of seamless interoperation of leading e-Infrastructures serving the

[13] See: www.ibbt.be/en/ilab/ilab-t.

[14] See: www.grid5000.fr.

European Research Area. The focus of the forum is the needs of the user communities that require services which can only be achieved by collaborating Infrastructures.

The membership of the forum is limited to representatives of large-scale, multi-national, multi-disciplinary Infrastructures. New members of the forum will be invited to join subject to agreement by a majority of the existing members. The initial membership was drawn from the following Infrastructures: EGEE, EGI, DEISA, PRACE, TERENA, GÉANT.

The forum will operate by consensus and will meet approximately once per quarter. Meetings will be hosted by members on a rotating basis, preferably associated with other major events. However the EEF website does not mention any activity or documents since mid- 2010.

European Intergovernmental Research Organization Forum (EIROforum);
web site: www.eiroforum.org

EIROforum is a collaboration between eight *European Intergovernmental scientific Research Organisations* (EIROs) that are responsible for infrastructures and laboratories: CERN, EFDA-JET, EMBL, ESA, ESO, ESRF, ILL and European XFEL. It is the mission of EIROforum to combine the resources, facilities and expertise of its member organisations to support European science in reaching its full potential.

3.8.2 Support Projects

We list in the following some support projects related to e-Infrastructures. We include in the list also the OSIRIS project.

The European eInfrastructures Observatory (e.nventory); web site: www.enventory.eu

The aim of the e.nventory project is to carry out a design study that will set the grounds towards the European eInfrastructures Observatory; a single-entry-point and one-stop-shop data warehouse, capable of representing multiple primary and convoluted indicators and benchmarks and a yardstick tool for progress monitoring and impact assessment of e-Infrastructures at regional and national level across the European Union and beyond. Through the collection and utilisation of appropriate indicators, the project will be able to monitor the status quo and evolution over time of e-Infrastructures development and communicate all findings to related stakeholders and to the public-at-large, in a seamless and impartial way.

Erina plus; web site: www.erinaplus.eu

ERINA+ aims to evaluate the impact of e-Infrastructure funded projects through the deployment of an effective socio-economic methodology. Already experimented within the ERINA study, the methodology is optimised and refined during the first phases of the project and then applied to a set of ongoing e-Infrastructure projects. The analysed projects are supported by the ERINA+ team in learning, understanding and applying the methodology which is then transferred to stakeholders and future project coordinators through communication activities. The Erina+ Model suggests four steps towards the socio-economic impact analysis of e-Infrastructure projects:

1) *Mapping the areas of impact* – it is important to have clear boundaries about what the analysis will cover, who will be involved in the process and how. Through engaging with the stakeholders an impact map will be developed, or a theory of change, which shows the relationship between inputs, outputs and outcomes.
2) *Defining the case of 'no action' (baseline) and alternatives* – this is the basis on which to compare improvements due to the project.
3) *Measuring or estimating impact indicators* for all options under (2) – this stage involves finding data to show whether outcomes have occurred and then valuing them. Having collected evidence on outcomes and monetised them, those aspects of change that would have happened anyway or are a result of other factors are eliminated from consideration.
4) *Exercise final analyses* – this stage involves adding up all the benefits, subtracting any negatives and comparing the result to the investment. Easily forgotten, this vital last step involves sharing findings with stakeholders and responding to them, embedding good outcomes processes and verification of the report.

OSIRIS Project; web site: www.osiris-online.eu

The main aim of the OSIRIS project initiative is to build the platform, mechanism and models required for securing the efficient involvement of Member States, Associated Countries and regions in developing cross border public-public partnerships and to establish a coordinated approach to future large scale investments in transnational European ICT RIs.

Expected results for OSIRIS project initiative are:

- the establishment of an Open living working group (platform) able to provide continuous analysis and recommendations on existing and future European ICT RIs regarding cross-border shared methodologies and best practices;
- the development of procedures, rules and management mechanisms to be codified in the Memorandum of Understanding for the continuation of the OSIRIS living working group for coordinated investments in large scale transnational ICT RI's in Europe;
- sustainability models and recommendations for future coordinated investments within and across European ICT RIs, with an emphasis on complementary or common planning of investments and investment policies;
- development of shared common cross-border methodologies and best practices to deploy and implement models for enhancing public-public partnerships that will leverage public-private ones.

Final Conclusion

The field of ICT Research Infrastructures is considerable and diversified, with widely varying collaboration models. In some cases there are functional *similarities* between them (for instance between network & DCIs), but there are also significant *differences* (for instance between the network integration model and the multi-facilities MNT model of collaboration).

Also, the *maturity* of collaboration models (integration & coordination, funding, users, industry involvement) is not comparable between different RIs, as in some domains there are already production-level infrastructures operational with well-defined governance (e.g., networks, DCIs), whilst in some other domains the picture is more fragmented (e.g., data infrastructures) or even still not defined (Future Internet).

A high-level list of seven relevant ICT Research Infrastructures has been surveyed in this document, outlining for each domain the relevant projects, the governance models and the challenges experienced. A finer granularity inventory of per-domain PA/NC-RIs collaboration models, listing thirty-one use cases is provided in Part II of this document.

Identified models and use cases will serve as a reference basis for the subsequent analysis that will be carried out by WP4.

Abbreviations and acronyms

RISGE-RG – Remote Instrumentation Services in Grid Environment Research Group (OGF)

ARI-WG – Access to Remote Instrumentation in a distributed environment Working Group (OGF)

ANNA – European Integrated Activity of Excellence and Networking for Nano and Micro-Electronics Analysis

ASCR – Academy of Sciences of the Czech Republic

BoD – Board of Directors

CERN – Conseil Européen pour la Recherche Nucléaire

CESNET – Czech academic network operator

CSEM – Centre Suisse d'Electronique et de Microtechnique

DEISA – Distributed European Infrastructure for Supercomputing Applications

DMTF – Distributed Management Task Force

EC – European Commission

ECDL – European Conference on Digital Libraries

ECRI – European Conferences on Research Infrastructures

EDGI – European Desktop Grid Initiative (DCI middleware provider / project)

EGEE – Enabling Grids for E-sciencE

EGI – European Grid Infrastructure

EGI_DS – EGI Design Study, project for the conceptual setup and operation of EGI

EEF – European E-Infrastructure Forum

EFII – European Future Internet Initiative

EIB – European Investment Bank

EIC – European Institute for innovation and Technology

EIRO – European Intergovernmental Research Organization (e.g., CERN, EMBL, ESA, …)

e-IRG – e-Infrastructure Reflection Group

EMI – European Middleware Initiative (DCI middleware provider / project)

ERA – European Research Area

ERC – European Research Council

ERIC – European Research Infrastructure Consortium

ESFRI – European Strategy Forum on Research Infrastructures

EU – European Union

EURAB – European Research Advisory Board

IaaS – Infrastructure as a Service

IE – Instrument Element

IPR – Intellectual Property Rights

IGE – Initiative for Globus in Europe (DCI middleware provider / project)

ISC – International Supercomputing Conference; European equivalent to the US based SC0x conference. Held annually in Germany.

FEDERICA – Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures

FIRE – Future Internet Research & Experimentation

FP6, FP7 – Framework Programmes 6, 7

GÉANT – Gigabit European Academic Network

HPC – High Performance Computing. Computing at the highest performance level at any given time; often used as a synonym for Supercomputing.

HET – High Performance Computing in Europe Taskforce. Taskforce by representatives from the European HPC community to shape the European HPC Research Infrastructure

ICC – Infrastructural Competence Centre

ICT – Information and Communication Technology

IE – Instrument Element

IMCS – Institute of Mathematics and Computer Science, University of Latvia

INFN – Istituto Nazionale di Fisica Nucleare

ISTAG – IST Advisory Group

KM3NeT – Kilometre Cube Neutrino Telescope

KIC – Knowledge and Innovation Community (see EIC)

LHC – Large Hadron Collider

MNT – Micro and Nano Technologies. In this deliverable it includes also photonics, organic electronics, etc. and similar activities in the hardware development

MNT-Europe – Staircase Towards European MNT Infrastructure Integration

MS – Member States

NGI – National Grid Initiative

NREN – National Research and Education Network

OLWG – Open Living Working Group

OSIRIS – towards an Open and Sustainable ICT Research Infrastructure Strategy

PPP – Private Public Partnership

PPC – Principal Partners Committee (for PRACE), also called hosting members

PRACE – Partnership for Advanced Computing in Europe

PRINS – Pan-European Research Infrastructure for Nano-Structures

RI – Research Infrastructure

RDI – Research Data Infrastructure

RO – Research Organization

RPF – Regional Partner Facilities

SDO – Standards Development Organisation

SME – Small and Medium sized Enterprise

SKA – Square Kilometre Array

Tier-0 (for EGI/LCG) – Denotes the first level of Grid systems (located at Cern). NGIs host the Tier-1 and Tier-2 systems.

Tier-0 (for HPC) – Denotes the apex of a conceptual pyramid of HPC systems. In this context the Supercomputing Research Infrastructure would host the Tier-0 systems. National or topical HPC centres would constitute Tier-1.

UNICORE – Uniform Interface to Computing Resources. Grid software for seamless access to distributed resources.

VRC – Virtual Research Community

VO – Virtual Organization

References:

[1] ARI-WG: http://www.ggf.org/gf/group_info/view.php?group=ari-wg;

[2] Bany Mohammed, A.A., Altmann, J., "A funding and governing model for achieving sustainable growth of computing e-infrastructures", Ann. Telecommun. (2010) 65:739–756;

[3] Berberich, F.,, Griffiths, E., "Report on options for a legal entity", PRACE Deliverable D2.1.1, 2008;

[4] Candiello, A., Cresti, D, Ferrari, T., Karagiannis, F. , Kranzlmueller, D. Louridas, P., Mazzucato, M., Matyska, L., Perini, L., Schauerhammer, K., Ullmann, K., Wilson, M. (2010) "A business model for the establishment of the, European Grid Infrastructure", CHEP 2009, J. Phys.: Conf. Ser. Vol.219 Part.06, 2011;

[5] COM(2008) 467, final 2008/0148 (CNS), Proposal for a COUNCIL REGULATION on the Community legal framework for a European Research Infrastructure (ERI) {SEC(2008) 2278} {SEC(2008) 2279} (presented by the Commission);

[6] Council conclusions of 20 November 2008 on the European digital library EUROPEANA (2008/C 319/07);

[7] Cresti, D., "The EGI Design Study Knowledge Base", EGI_DS Deliverable D2.2/D3.2/D3.5, December 2009, online: http://knowledge.eu-egi.eu/knowledge;

[8] Cruz-Castro, L., Sanz-Menéndez, L., Martínez, "Research centers in transition: patterns of convergence and diversity", J Technol Transf, April 2010;

[9] Dias, A.B., Douglas, S., "Final Report on the Peer Review Process", PRACE Deliverable D2.4.2, 2009;

[10] Dickmann, F., Brodhun, M., Falkner, J., Knoch, T.A., Sax, U., "Technology Transfer of Dynamic IT Outsourcing Requires Security Measures in SLAs", in: J. Altmann and O.F. Rana (Eds.): GECON 2010, LNCS 6296, pp. 1–15, 2010;

[11] DMTF, e-IRG Report on Data Management, 2009;

[12] Dzamila Bienkowska, Katarina Larsen and Sverker Sörlin, Public-private innovation: Mediating roles and ICT niches of industrial research institutes, 2010, Innovation: Management, Policy & Practice, (12), 2, 206-216;

[13] EEF: http://www.einfrastructure-forum.eu/;

[14] EGI.eu, Template for MoU between EGI.eu and VRCs, September 2010;

[15] Eickermann, T., "PRACE Final Report", PRACE Deliverable D1.5, 2010;

[16] e-IRG, "Blue Paper" 2010;

[17] e-IRG, "e-IRG White Paper 2011", v2.6, 5 April 2011;

[18] e-SciDR – Towards an e-Infrastructure for e-Science digital repositories - Final Report, 2008 for DG Information Society and Media Unit F – GÉANT and e-Infrastructure;

[19] e-SciDR Project Team, "Towards a European e-Infrastructure for e-Science Digital Repositories Final Report", DG Information Society and Media, Unit F – GÉANT and e-Infrastructure 2008;

[20] EARNEST, "Report on Organisation and Governance Issues", TERENA EARNEST, September 2007

[21] EGI Policy Group, "EGI/NGI Policy Session – Survey Analysis", regarding ERIC, presented at Vilnius, April 2011;

[22] EIROforum report, "Establishing New Research Infrastructures in Europe – The EIROforum Experience", March 2010;

[23] European Commission European Science Foundation, "Trends in European Research Infrastructures Analysis of data from the 2006/07 survey", European Commission, July 2007;

[24] EMI Project wiki, https://twiki.cern.ch/twiki/bin/view/EMI

[25] European Parliament, Directorate General For Internal Policies, "Evaluation of the European Research Area (ERA): Governance Aspects Study", IP/A/ITRE/ST/2008-13 PE 429.977, EN, November 2009 ;

[26] European Portal On Research Infrastructures' Services (online database): http://www.riportal.eu/public/index.cfm?fuseaction=ri.search;

[27] Future Internet Portal: www.future-internet.eu;

[28] Gentzsch, W., Girou, D., Kennedy, A., Lederer, H., Reetz, J., Riedel, M., Schott, A., Vanni, A., Vazquez, M., Wolfrat, J., "DEISA – Distributed European Infrastructure for Supercomputing Applications", J Grid Computing, March 2011;

[29] Gichoya D (2005) "Factors Affecting the Successful Implementation of ICT Projects in Government" The Electronic Journal of e-Government Volume 3 Issue 4, pp 175-184, available online at www.ejeg.com;

[30] Goldstein, H.A., Glaser, K., "Research universities as actors in the governance of local and regional development", J Technol Transf, September 2010;

[31] Griffiths, E., Subirada, F., Lichnewsky, A., "Report on analysis of adequate governance structure", PRACE Deliverable D2.2.1, 2008;

[32] Hey, T, Tansley, S. and Tolle, K, (Eds.) "The Fourth Paradigm: Data-Intensive Scientific Discovery." Microsoft Research. Redmond, Wash: 2009.;

[33] High level Expert Group on Scientific Data (HlEGSD) - Riding the wave - How Europe can gain from the rising tide of scientific data - Final report of the High Level Expert Group on Scientific Data – European Union, October 2010;

[34] Jeffery., K., "The Esfri Roadmap And Its Demands On The E-Infrastructure", Data Science Journal, Volume 9, July 24th 2010;

[35] Kamal, Mehruz and Qureshi, Sajda, "Classification of Countries Based on ICT Infastructure, Skills and Growth: The Role of ICT, Capacity in Emerging Economies" (2009). AMCIS 2009 Proceedings. Paper 426. http://aisel.aisnet.org/amcis2009/426;

[36] Kautta, M., Walshb, S.T., Bittnera, K., "Global distribution of micro–nano technology and fabrication centers: A portfolio analysis approach", Technological Forecasting and Social Change, Volume 74, Issue 9, November 2007, Pages 1697-1717;

[37] Knobloch, J. (Ed.), EGI_DS, "EGI Blueprint", EGI_DS D5.3, December 2008;

[38] Laaksonen, L., Järvinen, J., Koski, K., Rauhansalo, K., Polak, M., Kunszt, P., Griffiths, E., "Analysis of HPC Ecosystem", PRACE Deliverable D2.5.1, 2008;

[39] Lucas, M., Sands, A., Wolfe, D.A:, "Regional Clusters in a Global Industry: ICT Clusters in Canada", European Planning Studies Vol. 17, No. 2, February 2009;

[40] Mazzetti, P., "Research Strategies for the development of a Civil Protection E-Infrastructure", CYCLOPS Deliverable D16, November 2008;

[41] Menz, N., Ott, I., "The Geography of Nanotechnologies: The role of Specialization and Diversification",June 23th 2010 (found on the web);

[42] Mowery, D.C., "Nanotechnology and the US national innovation system: continuity and change", J Technol Transf, ,February 2011;

[43] OECD report, "Establishing Large International Research Infrastructures: Issues and Options", December 2010;

[44] Oliveira, J., "Project Results Presentation", CYCLOPS Deliverable D18, November 2008;

[45] Palmberg, C., H. Dernis and C. Miguet (2009), "Nanotechnology: An Overview Based on Indicators and Statistics", OECD Science, Technology and Industry Working Papers, 2009/7, OECD Publishing;

[46] Perini, L. (Ed), "D3.2 Final EGI Functions Definition", EGI_DS Deliverable D3.2, May 2009;

[47] RISGE-RG: http://www.ogf.org/gf/group_info/charter.php?review&group=risge-rg;

[48] Rizzuto, C., "Research Infrastructures and the Europe 2020 Strategy", Science Business, 2010;

[49] Rizzuto, C., "European Roadmap for Research Infrastructures – Implementation Report 2009", ESFRI, 2009;

[50] Sargent Jr., J. F., "Nanotechnology: A Policy Primer", Congressional Research Service, 7-5700, RL34511, Jan 4th, 2010;

[51] Shapira, P., Youtie, J., Kay, L., "National innovation systems and the globalization of nanotechnology innovation", J Technol Transf, February 2011;

[52] Sorani, R., "Perspective on cooperation with existing projects and initiatives", CYCLOPS Deliverable D3. August 2006;

[53] Suvagnargues-Lesage, S., "Business Process Analysis Document". CYCLOPS Deliverable D6, August 2007;

[54] Suvagnargues-Lesage, S., "Existing analysis document". CYCLOPS Deliverable D8, August 2007;

[55] TERENA wiki, https://wiki.terena.org;

[56] Thursby, J., Thursby, M., "University-industry linkages in nanotechnology and biotechnology: evidence on collaborative patterns for new methods of inventing", J Technol Transf, February 2011;

[57] Weltring, Klaus-Michael (Ed), EuroBioNano Final Report "Concept for a European Infrastructure in Nanobiotechnology", January 2010;

[58] Working Group on Future Internet (WGFI), "The Future of Internet, Report from the National ICT Research Directors WGFI", European Commission Information Society & Media, November 2008.